

Visão Computacional

Autor: Hemerson Pistori

VERSÃO PRELIMINAR DE APOSTILA DESENVOLVIDA PARA O
CURSO DE ESPECIALIZAÇÃO EM BIOTECNOLOGIA
Parceria Universidade Católica Dom Bosco e Portal Educação

Sumário

UNIDADE 1 - Introdução.....	3
1. 1 História, Abrangência e Inter-relações.....	4
1. 2 Exemplos de aplicações.....	9
1. 3 Ferramentas.....	14
1. 4 Literatura complementar.....	15
UNIDADE 2 - Formação da imagem, captura e representação.....	17
2.1 Visão humana e dispositivos de captura.....	17
2.2 Calibração.....	21
2.3 Ruído.....	22
UNIDADE 3 - Pré-processamento e Filtros.....	24
3.1 Histogramas.....	27
3.2 Suavização.....	28
3.3 Realce e Detecção de bordas.....	31
3.4 Morfologia Matemática.....	32
UNIDADE 4 – Segmentação.....	33
4.1 Segmentação por limiarização.....	34
4.2 Segmentação por agrupamento.....	36
4.3 Segmentação por ajuste de modelos.....	39
UNIDADE 5 – Extração de atributos.....	42
5.1 Cores.....	42
5.2 Forma.....	44
5.3 Textura.....	45
5.4 Pontos de interesse.....	45
5.5 Seleção e Redução de atributos.....	48
UNIDADE 6 - Reconhecimento e Detecção.....	50
6.1 Casamento de modelos.....	50
6.2 Aprendizagem automática.....	52
6.3 Métodos sintáticos ou estruturados.....	54
UNIDADE 7 – Rastreamento.....	57
7.1 Fluxo Óptico.....	57
7.2 Filtros preditivos.....	58

UNIDADE 1 - Introdução

A grande maioria dos textos sobre visão computacional disponíveis atualmente é voltada para estudantes ou profissionais das áreas da computação, engenharias ou afins e requerem conhecimento prévio razoável em temas como álgebra, cálculo, estatística e programação. Como esta apostila tem como alvo um público com formação bastante diversificada, buscou-se apresentar a área da visão computacional de forma abrangente e leve, sem entrar em detalhes envolvendo programação de computadores ou fundamentação matemática, e concentrando-se em aspectos de caráter mais prático e intuitivo. Espera-se que o leitor, através deste texto, possa ter uma noção geral do que é a visão computacional, como ela pode ser aplicada, quais os principais problemas e técnicas utilizadas, quais são os conceitos fundamentais e a terminologia utilizada. É um texto para estudantes e profissionais que não têm, necessariamente, como objetivo principal se tornar um desenvolvedor de sistemas de visão computacional, mas que desejam ter uma boa noção do que pode ser feito com este tipo de sistema e além disso, que possam interagir produtivamente com engenheiros e cientistas da computação especialistas em visão computacional para ajudá-los no desenvolvimento deste tipo de tecnologia. Mesmo para aqueles interessados em se tornar um especialista em visão computacional, o



Figura 1: Quadro de um vídeo que apresenta as etapas de um sistema de visão computacional. Link: <http://youtu.be/1b781K1Cghg>

texto pode servir como uma primeira introdução mais “amigável” ao assunto. Ao final deste capítulo são listados alguns livros que podem ser consultados para um aprofundamento nos temas abordados nesta apostila. Também está disponível no link associado à Figura 1, um vídeo que introduz alguns dos principais tópicos estudados em visão computacional e que formam as unidades desta apostila.

1. 1 História, Abrangência e Inter-relações

O termo visão computacional é relativamente novo mas os conceitos estudados e as técnicas desenvolvidas nessa área da computação começaram a surgir a mais de meio século, juntamente com as áreas mais “tradicionais” como inteligência artificial e processamento digital de imagens. Ainda hoje, alguns autores apresentam a visão computacional como uma sub-área da inteligência artificial e outros como parte do processamento digital de imagens, mas a tendência parece ser a consolidação da visão computacional como uma área própria, com uma importante intersecção com essas duas outras áreas.

O objeto principal da visão computacional é a imagem digital, que pode ser hoje em dia obtida através de uma infinidade de sensores e equipamentos, desde as quase onipresentes câmeras fotográficas digitais, filmadoras, *smartphones*, *webcams* e câmeras de segurança até microscópios, telescópios, satélites, tomógrafos e aparelhos de ultrassom e de ressonância magnética. É justamente a grande disponibilidade e variedade de aparelhos capazes de gerar imagens digitais que fez com a área da visão computacional ganhasse um grande destaque como campo de estudo e de desenvolvimento de novos produtos nos últimos anos. A quantidade de imagens geradas diariamente nos dias de hoje é imensa e na maioria dos casos, obter informações relevantes a partir dessas imagens através da inspeção visual, mesmo que seja apenas para classificá-las e organizá-las, tornou-se uma tarefa hercúlea.

A visão computacional é justamente a área da computação que visa a criação de programas de computador capazes de extrair de forma automática ou ao menos semi-automática informações relevantes contidas em imagens digitais para resolver ou auxiliar na solução de problemas que dependam direta ou indiretamente dessas imagens. Pode-se também dizer que a visão computacional é a área da computação

que busca reproduzir nos computadores o sentido da visão dos seres humanos ou ao menos resolver problemas que para os seres humanos envolveriam a utilização da visão. Embora algumas das técnicas da visão computacional tenham inspiração no conhecimento ainda incompleto que já temos sobre o funcionamento da visão humana, geralmente os engenheiros e cientistas que pesquisam nessa área não se limitam a buscar entender primeiro a visão humana para então tentar reproduzi-la. Muitas das técnicas empregadas atualmente surgiram sem qualquer embasamento na biologia humana ou animal de forma geral, e mesmo assim conseguem resolver de forma eficiente alguns problemas específicos.

Como o sistema de visão humano é extremamente sofisticado, com processos que ocorrem de forma inconsciente e bastante eficiente, as pessoas que não conhecem a área da visão computacional acabam subestimando as dificuldades inerentes a diversas tarefas que envolvam a interpretação ou extração de informações a partir de imagens. Como para o ser humano é muito fácil, por exemplo, reconhecer objetos em imagens, assumi-se que trata-se de um problema simples de ser resolvido por computadores, no entanto, embora já tenhamos construído computadores capazes de ganhar de um campeão mundial de xadrez, ainda estamos muito longe de ter um computador capaz de se aproximar da capacidade de uma criança de 4 anos de idade de distinguir diferentes objetos presentes em imagens corriqueiras, com utensílios domésticos, móveis e pessoas, por exemplo. Mesmo estando longe de imitar a grande capacidade e flexibilidade do ser humano no uso da visão para resolver problemas, os avanços ocorridos nas últimas décadas, na área da visão computacional, já nos permitem solucionar diversos problemas relevantes, principalmente quando podemos restringir o escopo das imagens que irão alimentar o sistema computacional e existe um grande otimismo em relação aos tipos de problemas que poderão ser resolvidos nos próximos anos a partir de novos avanços nesta área.

A relação da visão computacional com o processamento digital de imagens (PDI) é bastante estreita. Em processamento digital de imagens desenvolve-se técnicas de manipulação de imagens que podem ser muito úteis como pré-processamento para os módulos do sistema de visão computacional que realizam a extração da informação. As unidades 2, 3, 4 e 5 desta apostila tratam de assuntos que pertencem tanto à área da visão computacional quanto ao processamento digital

de imagens. Uma maneira de se distinguir estas duas áreas é que, embora nas duas se tenha como entrada principal do sistema uma imagem, no processamento digital de imagens temos como saída uma outra imagem (E.g: com melhor contraste, mais clara ou apenas com as bordas dos objetos), enquanto na visão computacional temos informações em outros formatos, como descrições textuais, medições de áreas, contagens de objetos, relatórios, estatísticas ou até mesmo uma tomada de decisão, como é o caso, por exemplo, dos sistemas de visão ligados a alarmes e robôs que podem responder ou agir automaticamente com base em informações capturadas por câmeras.

Existe uma área chamada visão de máquina que se distingue da visão computacional por envolver o desenvolvimento de equipamentos específicos, geralmente voltados para a indústria e com softwares embutidos¹. Já na visão computacional, de modo geral, temos o desenvolvimento de softwares que funcionam em computadores de uso geral como *laptops*, *desktops* e *smartphones*. Os fundamentos das duas áreas, no entanto, são basicamente os mesmos.

Uma outra área que se relaciona bastante com a visão computacional é o reconhecimento de padrões que trata do desenvolvimento de programas de computador e equipamentos capazes de classificar ou identificar coisas como o estilo de uma música que está sendo tocada, o idioma que está sendo falado, o estilo literário de um texto, a possibilidade de um determinado e-mail ser *spam* ou não, a identidade de um criminoso a partir de fragmentos do DNA, a probabilidade de um terremoto com base em sinais sísmicos e o humor de uma pessoa analisando imagens da sua face. É possível notar por este pequeno conjunto de exemplos que as aplicações para o reconhecimento de padrões são muito amplas e o tipo de dados ou sinais envolvidos variam desde textos, sequências de caracteres, sons, sinais sísmicos e evidentemente, imagens. Como uma parte significativa do que se faz em visão computacional envolve a identificação de coisas e eventos em imagens, as técnicas de reconhecimento de padrões são bastante úteis.

Uma das maneiras de se realizar reconhecimento de padrões é através das técnicas desenvolvidas em uma outra área da computação chamada de

¹Diferentemente do software comum, que pode ser instalado e executado em computadores de uso geral, o software embutido fica dentro de um equipamento que não é um computador (E.g: automóveis, geladeiras, televisores, etc).

aprendizagem de máquina, que busca a criação de computadores e outros equipamentos que sejam capazes de “aprender sozinhos” a resolver problemas ou realizar tarefas sem que um programador tenha que programá-los explicitamente. Na aprendizagem de máquina, o usuário precisa “apenas” fornecer exemplos daquilo que precisa ser aprendido e o software é capaz de abstrair a partir desses exemplos um modelo a ser utilizado pelo próprio software para resolver problemas similares, apresentados posteriormente. A aprendizagem automática ainda é uma área em franco desenvolvimento e estamos distantes de programas capazes de emular a flexibilidade e eficácia que o ser humano possui para aprender a partir de suas experiências. No entanto, em contextos específicos já existem diversos casos de sucesso, sendo um dos mais conhecidos os programas de detecção de *spam* capazes de melhorar seu desempenho conforme o usuário vai assinalando se determinada mensagem trata-se ou não de um *spam*. Quando a aprendizagem de máquina é aplicada em problemas envolvendo análise de imagens, temos a intersecção entre aprendizagem de máquina e visão computacional.

Para alguns autores, a inteligência artificial engloba todas as áreas citadas até agora, incluindo a visão computacional, uma vez que foi criada inicialmente com o objetivo bastante ambicioso de reproduzir em uma máquina aquilo que nos seres humanos chamamos de inteligência. A inteligência humana permite-nos interpretar e utilizar de forma eficaz informações obtidas através dos nossos sentidos, que incluem a visão, que opera justamente em imagens e portanto, temos uma relação óbvia entre visão humana e visão computacional. Aliás, algumas das técnicas da visão computacional foram inspiradas no que já se conhece sobre o funcionamento da visão humana. Conforme a área da inteligência artificial foi se aprofundando nos problemas relacionados com a construção de sistemas inteligentes, mesmo que de forma limitada e especializada, algumas das suas sub-áreas acabaram se tornando um novo campo de estudo *per se*, como foi o caso da visão computacional. No entanto, as relações entre essas áreas continuam sendo muito importantes e ricas.

Por fim, uma outra área da computação que se relaciona com a visão computacional é a computação gráfica, que também tem como objeto central a imagem, mas enquanto na visão computacional a imagem é a entrada para o sistema, na computação gráfica ela é a saída. Ou seja, na primeira, partimos das imagens para chegar às informações em outros formatos, enquanto na computação

gráfica temos a informação sobre o que queremos apresentar ou representar e o resultado final é uma imagem ou um vídeo. Tratam-se de problemas bastante distintos, com técnicas e fundamentos distintos, embora existam alguns problemas em que computação gráfica e visão computacional são utilizados de forma complementar, como por exemplo, na construção de sistemas de realidade aumentada, em que se busca combinar uma imagem de cenas ou objetos reais com imagens geradas por computador. O *google glass*® é talvez um dos dispositivos mais conhecidos atualmente para produzir realidade aumentada. Outra área em que computação gráfica e visão computacional atuam complementarmente é na geração de animações para o cinema em que os movimentos dos personagens animados são produzidos a partir da análise de vídeos de pessoas reais realizando os mesmos movimentos. Na Figura 2 temos um quadro de um vídeo apresentando algumas das relações da visão computacional com outras áreas afins.

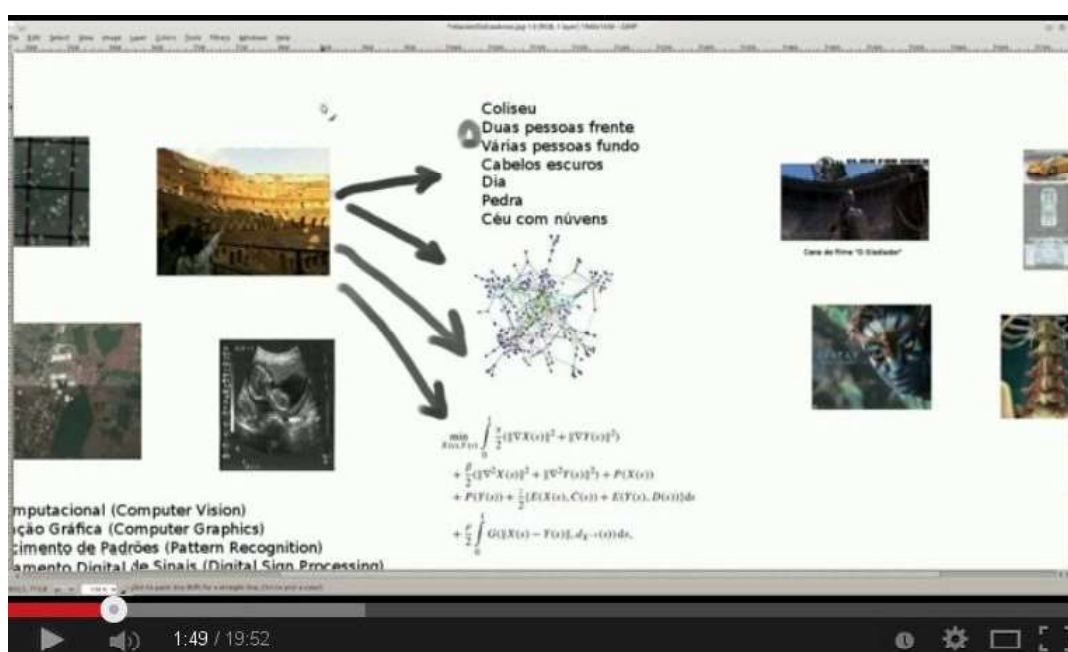


Figura 2: Vídeo no youtube apresentando a relação da visão computacional com outras áreas de conhecimento afins. Fonte: <http://youtu.be/CKO3g7ziX0I>

Exemplos de Aplicações de Visão Computacional



Figura 3: Playlist do Prof. Pistori com uma seleção de vídeos mostrando aplicações de visão computacional: <http://www.youtube.com/playlist?list=PL42BEFB9AC2908DBA>

1. 2 Exemplos de aplicações

Uma vez que uma quantidade enorme de tarefas e ações realizadas pelo ser humano depende da visão, o potencial para aplicação da visão computacional é imenso. Com os avanços constantes tanto na capacidade de processamento dos computadores, quanto na disponibilidade de dispositivos de captura de imagens e também de técnicas cada vez mais sofisticadas, as barreiras tecnológicas e econômicas também vão sendo quebradas e novos sistemas vão surgindo com frequência cada vez maior. Já existem inúmeras empresas especializadas no desenvolvimento de soluções usando visão computacional e a lista de áreas em que este tipo de tecnologia já vêm sendo aplicada aumenta a cada dia. Na Figura 3 temos um link para uma coleção de vídeos (*playlist*) de aplicações de visão computacional em áreas como agricultura, meteorologia, medicina, segurança, biologia, astronomia, ecologia e marketing.

Na biotecnologia, a visão computacional pode ser aplicada desde a prospecção de biodiversidade até os testes finais de produtos biotecnológicos, sempre que o processo envolver algum tipo de análise visual de imagens. A seguir, apresentaremos algumas das aplicações que foram ou estão sendo desenvolvidas pelo grupo de pesquisa em visão computacional da

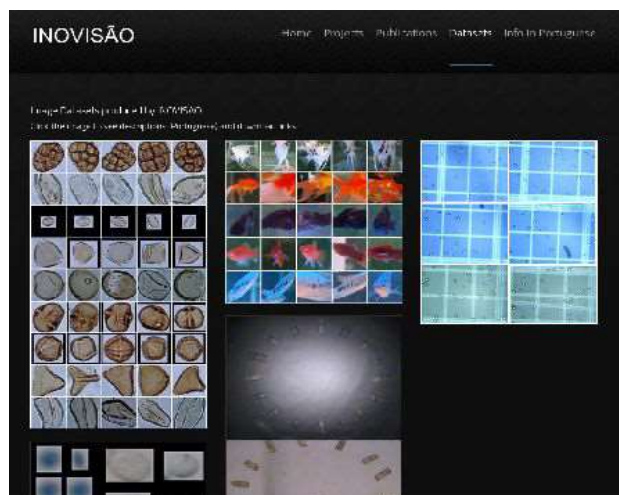


Figura 4: Site do INOVISAO: www.gpec.ucdb.br/inovisao

UCDB que está ligado aos programas de pós-graduação *stricto sensu* em biotecnologia e também em ciências ambientais e sustentabilidade agropecuária da mesma universidade. A Figura 4 mostra uma das telas do site do INOVISAO que disponibiliza acesso a imagens utilizadas nos experimentos realizados pelo grupo.

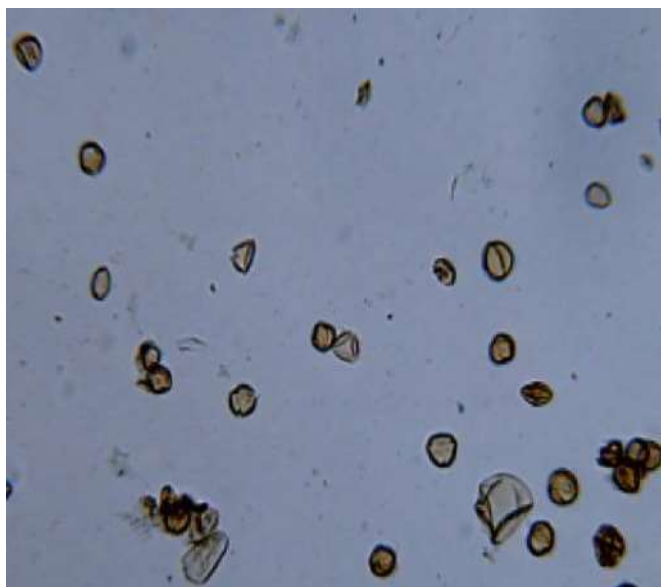


Figura 5: Grãos de pólen presentes em uma imagem microscópica de uma amostra de mel

O primeiro sistema de visão computacional que iremos apresentar tem como objetivo auxiliar na identificação da origem geográfica de uma amostra de mel. A partir de uma imagem obtida através de um microscópio, como a que aparece na Figura 5, o sistema deve automaticamente identificar as espécies de pólen presentes e calcular as proporções de cada espécie. A partir dessa informação, utilizando dados sobre as características da flora de determinadas regiões, o sistema irá oferecer ao usuário estimativas sobre a localização em que o mel foi produzido. Esse tipo de sistema pode ser utilizado para identificar fraudes na comercialização do mel e permitir a implementação de certificação de origem controlada para mel de alta qualidade produzido em algumas regiões Brasileiras.

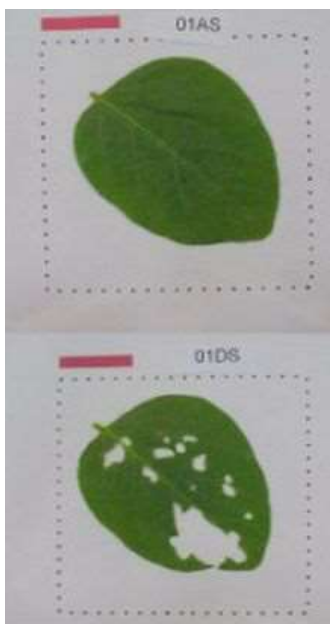


Figura 6: Folha de soja antes e depois da herbivoria

Em um outro sistema de visão computacional desenvolvido pelo INOVISAO pode-se realizar automaticamente a medição do efeito de herbivoria em diferentes espécies transgênicas de folhas de soja. A partir de imagens como a da Figura 6, o software gera um relatório contendo informações sobre o percentual da folha que foi, por exemplo, atacado por lagartas. O grupo está trabalhando agora no desenvolvimento de sistemas de visão computacional para analisar imagens capturadas através de veículos aéreos não tripulados, os chamados VANTs, DRONES ou UAVs), e realizar medições e identificação de eventuais problemas diretamente no campo. Com os preços dos VANTs e das câmeras que podem ser acopladas a esses VANTs caindo anualmente, existe a previsão de que

no futuro próximo teremos o início da utilização em larga escala desse tipo de veículo para agricultura de precisão. Como os VANTs conseguem sobrevoar uma plantação a poucos metros de distâncias, câmeras de alta resolução, como as que encontramos hoje com facilidade em qualquer mercado, são capazes de obter imagens até de pequenos insetos presentes na plantação. Sistemas de visão computacional serão essenciais, no entanto, para tornar esse tipo de aplicação viável, uma vez que a quantidade de imagens geradas é muito grande para ser analisada por seres humanos.

Em um terceiro projeto do INOVISAO, está sendo desenvolvido um sistema para auxiliar na análise de diferentes estratégias para



Figura 7: Arena para experimentos com carunchos e amostras de bambu

tornar o bambu menos vulnerável ao caruncho, um inseto que prejudica a utilização da planta e precisa ser combatido. O sistema deverá ser capaz de rastrear carunchos colocados em uma arena em que estão dispostas amostras de bambu com diferentes tratamentos ou de diferentes espécies. Uma filmadora colocada acima da arena captura imagens que são depois processadas pelo sistema de visão computacional que oferecerá informações como trajetórias, velocidades e distribuição dos carunchos durante os experimentos. Informações desse tipo podem ser úteis na análise do comportamento dos carunchos frente a determinada estratégia para afastá-lo da planta. A Figura 7 ilustra um tipo de imagem que pode servir de entrada para o sistema de visão computacional.

Entre os outros sistemas desenvolvidos ou em desenvolvimento pelo INOVISAO temos o LARVIC, que calcula a taxa de mortalidade da larva do *Aedes aegypt* em experimentos com larvicidas, o TOPOLINO, para análise de comportamento de camundongos em experimentos como labirinto em cruz, *watermaze* e campo aberto, o CORIU, que analisa imagens de couro bovino para identificar e calcular a área de defeitos como marca a ferro, furo de esfolia, sarnas, entre outros, o FISHCV, para identificação de espécies de peixe, o ESTRIVIC, para auxílio na análise da evolução do paciente em tratamentos de estrias e o BIOVIC, que realiza identificação e contagem de leveduras viáveis e inviáveis em amostras do mosto de cana-de-açúcar para auxiliar no controle microbiologia em usinas de produção de etanol. A Figura 8 apresenta imagens do TOPOLINO e do CORIU.

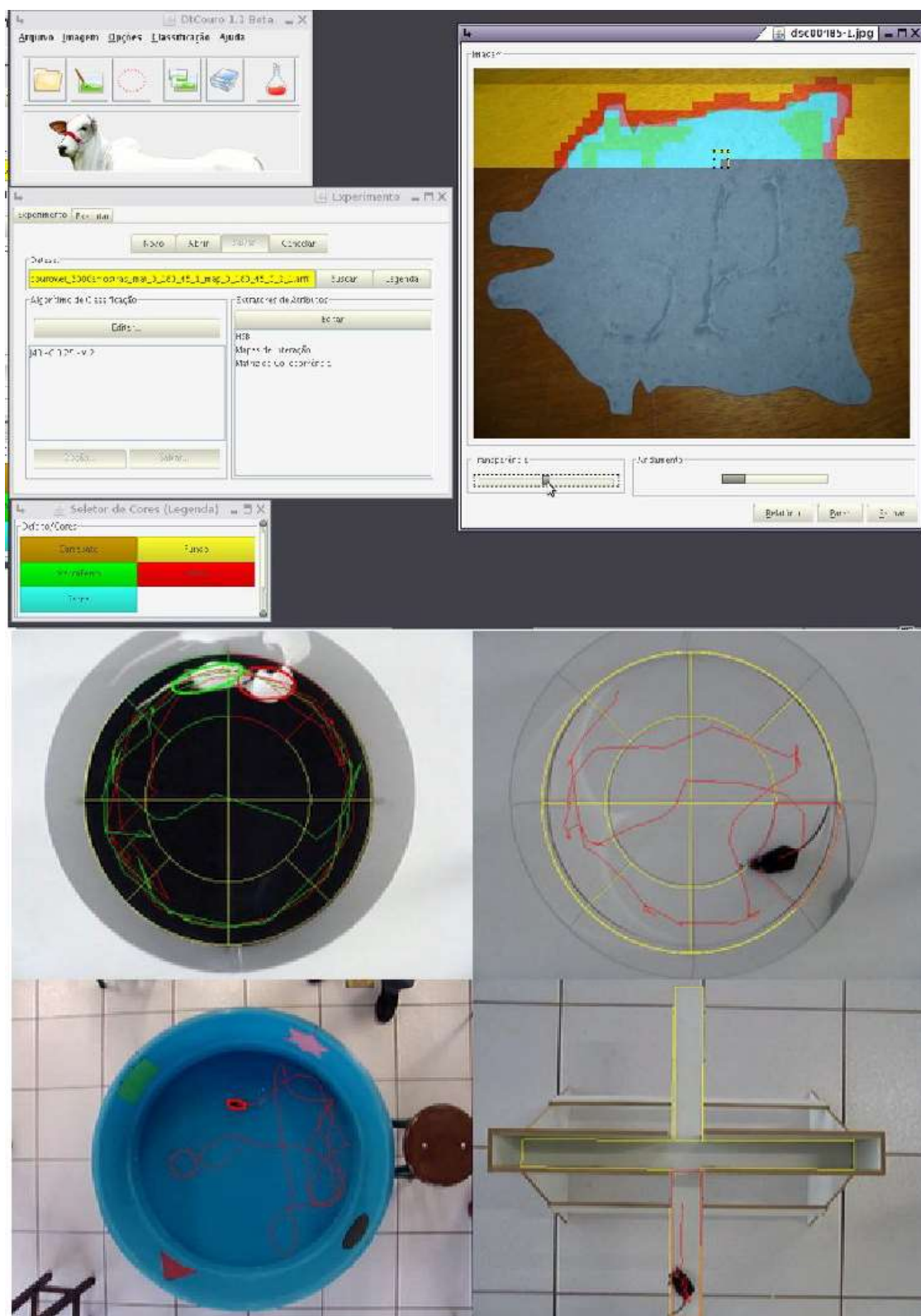


Figura 8: Telas dos softwares CORIU, de detecção de defeitos em couro bovino, e TOPOLINO, de análise de experimentos com camundongos

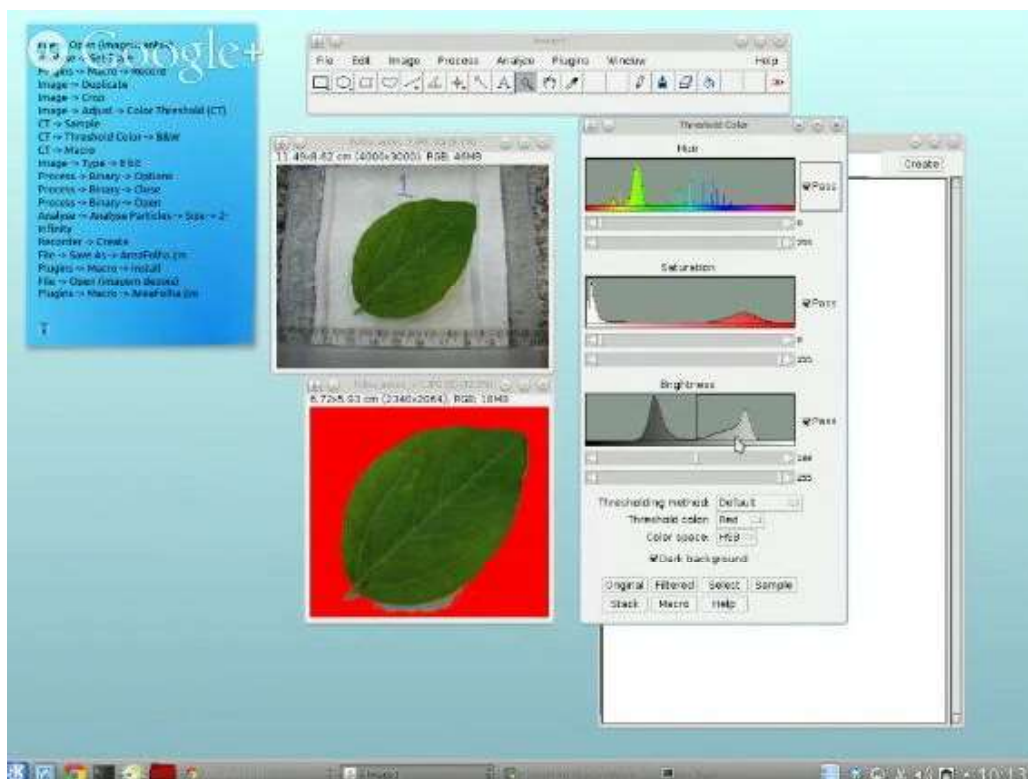


Figura 9: Tela de um vídeo sobre o ImageJ sendo aplicado na solução de um problema real. Link: <http://www.youtube.com/watch?v=1L49XO6NrSY>

1. 3 Ferramentas

Existem diversas ferramentas computacionais para auxiliar programadores no desenvolvimento de sistemas de visão computacional. Neste texto, iremos nos concentrar em uma que é bastante conhecida mundialmente na área e que possui um módulo que pode ser utilizado por usuários leigos no assunto para explorar alguns aspectos das técnicas de visão computacional e aprofundar seu conhecimento através da experimentação direta. Trata-se de um software gratuito chamado ImageJ² que pode ser obtido na Internet através do link que aparece na Figura 9, essa Figura mostra um dos quadros de um vídeo que introduz o ImageJ através de um exemplo. Como vários dos conceitos apresentados nesta apostila podem ser explorados através do ImageJ, recomendamos que o leitor tente instalar e aprender a usar o software consultando as diversas apostilas disponíveis na

² O software ImageJ pode ser obtido neste endereço: <http://rsbweb.nih.gov/ij/>

Internet. Para quem sabe programar, o ImageJ pode ser utilizado como uma biblioteca em Java. Também existe um outro software gratuito, chamado GIMP³, que foi utilizado para construir alguns dos exemplos desta apostila é que pode ajudar na compreensão de alguns conceitos, principalmente aqueles relacionados com os assuntos das unidades 2 e 3. Uma alternativa para trabalhar com visão computacional em C ou C++ é a biblioteca OpenCV⁴, uma das mais utilizadas na atualidade, e que possui recursos interessantes para desenvolvimento de aplicativos de visão computacional para celulares, *smartphones* e *tablets*. Existem também interfaces que permitem a utilização da OpenCV em Java e Python. Por fim, para utilização de técnicas de aprendizagem automática nas aplicações de visão computacional, recomendamos a biblioteca utilizada pelo grupo INOVISAO que é chamada Weka⁵. Tanto o OpenCV quanto o Weka são gratuitos e com códigos-fonte abertos.

1. 4 Literatura complementar

A seguir são listadas algumas referências para os principais livros-texto utilizados em cursos avançados de visão computacional. Alguns desses livros, como os dois primeiros, foram disponibilizados gratuitamente pelos seus autores e podem ser encontrados facilmente na Internet

DAVIES, E. R. **Computer and Machine Vision**: Theory, Algorithms, Practicalities. 4a Ed. Elsevier, 2012.

SZELISKI, R. **Computer Vision**: Algorithms and Applications. Springer, 2010.

PRINCE, S. J. D. **Computer Vision**: Models, Learning and Inference. Cambridge University Press, 2012.

FORSYTH, David and PONCE, Jean. **Computer Vision**: A Modern Approach. 2a Ed. Prentice Hall, 2011.

NIXON, Mark S. and AGUADO, Alberto S. **Feature Extraction & Image Processing**. 2a Ed. Newnes, 2008.

³ O GIMP pode ser obtido aqui: <http://www.gimp.org/>

⁴ O site do OpenCV é este: <http://opencv.org/>

⁵ O weka está disponível para download em <http://www.cs.waikato.ac.nz/ml/weka/>

GONZALEZ ,R.C., WOODS , R.E. **Digital Image Processing**. 3a. Ed. Edgard Blucher, 2007

DUDA, R.O., HART, P. E. and STORK, D. G. **Pattern Classification**. Wiley, 2000.

BRADSKI, G.; KAEHLER, A. **Learning OpenCV: Computer Vision in C++ with the OpenCV Library**. O'Reilly, 2012.

WU, Q.; MERCHANT, F.; CASTEMAN, K. **Microscope Image Processing**. Academic Press, 2008.

FLACH, P. Machine Learning: **The Art and Science of Algorithms that Make Sense of Data**. Cambridge, 2012.

BALLARD, D. H.; BROWN, C. M. **Computer Vision**. Prentice Hall, 1982.

JAIN, R.; KASTURI, R.; SCHUNCK, B. G. **Machine Vision**. McGraw-Hill, 1995.