



UNIVERSIDADE  
FEDERAL DO CEARÁ



CELL RECOGNITION  
FOR INSPECTION OF CERVIX

# CRIC HACKATHON 2017

COMPUTER VISION, BIG DATA & MACHINE LEARNING

Dani Ushizima  
Romuere Silva  
Flávio Araújo



# Objetivo

- O hackathon visa possibilitar que os participantes desenvolvam trabalhos de visão computacional, visualização, aprendizado de máquina e bases de dados gigantes;
- Cada equipe deve ser composta por até 3 membros (sem exceção) e terá tema diferente, então não se preocupe - não tem como colar!
- Teremos snacks durante o evento;
- Premiação em dinheiro (valor a ser divulgado) ao melhor (1o. lugar) projeto.

# Temas

1. Investigação de epidemias - desenhar ferramentas para metadado e geolocalização em mapas usando séries temporais (DB: mosquito outbreak);
2. Busca otimizada por similaridade;
3. Cortador automático de células;
4. Métodos de detecção de saliência.

# Hackathon

- Terá duração de 24 horas, sem interrupção;
- Os concorrentes poderão utilizar as salas reservadas para o evento CRIC 2017 para desenvolvimento do projeto;
- Alimentos e bebidas serão doados pelos organizadores do evento.

# Ajuda externa ao time

- Três professores estarão disponíveis durante o evento para ajudar com possíveis dúvidas, com agendamento de horários a ser divulgado.

**\*\* ATENÇÃO:** Todos os códigos submetidos serão publicados no site do evento bem como no GITHUB do projeto.

# Apresentação

- Cada equipe terá a chance de mostrar o que conseguiu realizar através de uma concisa apresentação de 10 min;
- Será realizada na terça-feira (04/07/2017) às 10:00 horas;
- Junto com a apresentação o time deverá entregar o código-fonte gerado.

Obs: apresentações que passem de 10 minutos serão penalizadas 1 ponto/minuto, que serão subtraídos da nota final.

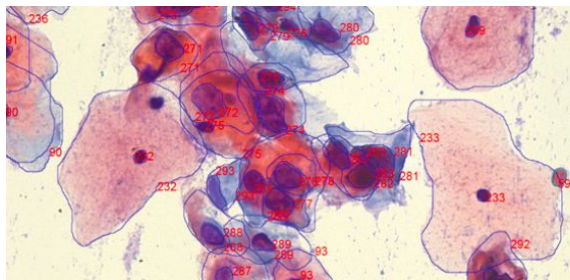
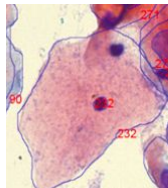
# CrITÉrios de AvaliaÇão

- Linguagem de Programação (python e/ou free software preferível);
- Criatividade;
- Precisão;
- Apresentação;
- Integração do Time;
- Engenharia de Software.

# Payattention

Encontre a célula mais semelhante

?





# Busca otimizada por similaridade

## Objetivo

1. Encontrar uma amostra através de seu respectivo vetor de características;
2. Busca otimizada imaginando que o banco de imagens terá milhões de amostras até o final de 2017;
3. Dada uma amostra, encontrar as 10 outras amostras mais semelhantes
4. Utilizar persistência de dados. Ex: SQL.

## Dados

1. Há 2 planilhas num arquivo csv:
  - a. cada amostra corresponde a uma linha e cada coluna corresponde a uma medida da amostra
  - b. os labels de cada amostra (0-5)
2. Amostras pertencem ao [cricdatabase.com.br](http://cricdatabase.com.br)

Facilitamos seu trabalho: encontre todas as medidas em:  
<https://www.dropbox.com/s/nb7qg7h4g3zz4r5/busca%20hackathon.zip?dl=0>

# Hadouken

Entenda a rota do mosquito *Aedes sp.*



# Estatísticas e visualização de rotas

## Objetivo

1. O objetivo é desenhar um protótipo de ferramenta que auxilia os investigadores a visualizar rotas utilizadas pelo mosquito e possíveis novos pontos baseados em direções anteriores. Por exemplo, dado que existe uma epidemia que começa em Catuana, depois vai a Guararu, então Caucaia, qual a probabilidade de as pessoas de Fortaleza serem as próximas vítimas?

## Dados

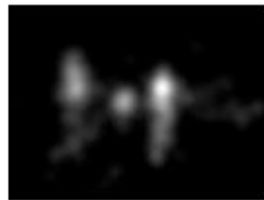
1. Há 2 planilhas num arquivo excel:
  - a. cidades brasileiras e georef
  - b. lista de ocorrências de doença relacionada ao mosquito
2. Dados são parcialmente fictícios e são provenientes de:  
<https://www.binapratika.com.br/port-al-transparencia-odv>
3. Caso queira ampliar o volume de dados, divirta-se.

Facilitamos seu trabalho: encontre as planilhas em:

<https://www.dropbox.com/s/rq8oyi5jxc0ezcf/MUNICIPIOS-LATITUDE%20E%20LONGITUDE-BRASIL.xlsx?dl=0>

# Scorpius

Quais pontos chamam atenção?



# Métodos de detecção de saliência

## Objetivo

1. Entender como relacionar mapas de saliência com dados de observadores rastreados com “eye tracker”
2. Testar 2+ métodos de detecção de saliência dada a base do MIT
3. Ilustrar com uma aplicação de detecção de saliência

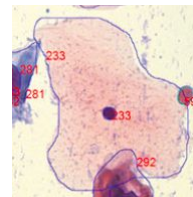
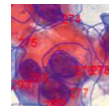
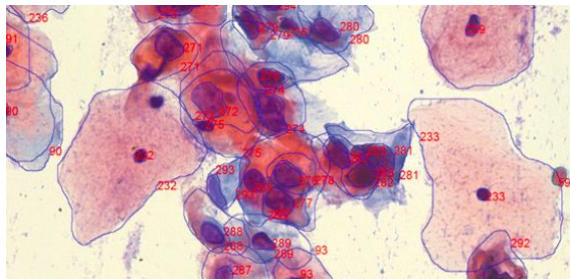
## Dados

1. 300 imagens;
2. dados de 39 observadores rastreados com “eye tracker”
3. [saliency.mit.edu/results\\_mit300.html](http://saliency.mit.edu/results_mit300.html)

Facilitamos seu trabalho: encontre imagens e realidade terrestre em:  
[saliency.mit.edu/results\\_mit300.html](http://saliency.mit.edu/results_mit300.html)

# PaoDeQueijo

# Cortador automático de células



# Detector automático de células

## Objetivo

1. Encontrar o ponto central de núcleos de células cervicais para cortar uma janela que contém a célula;
2. Detecção automática, **sem** interação manual, isto é, na mão;
3. Dada uma imagem de campo de lâmina de Papanicolau, salve as coordenadas daquela imagem.

## Dados

1. 30 imagens com cerca de 20 células em cada;
2. 22 imagens serão disponibilizadas para treino e 8 serão utilizadas para avaliação;
3. Link da base de treino:  
[https://www.dropbox.com/s/r2jixnm7qrun1s/find\\_cells\\_hackathon.zip?dl=0](https://www.dropbox.com/s/r2jixnm7qrun1s/find_cells_hackathon.zip?dl=0)

# Divirtam-se e boa sorte!

