
Guia do usuário: Viabilidade celular

Kayllany L. S. Oliveira, Daniel C. Vieira, José G. C. Pereira, João V. S. Guerra

November 29, 2024

Table of Contents

1	Introdução	3
2	Requisitos	3
3	Acessando o HPCC Marvin	3
4	Preparando o Repositório ParasiteCoLocalization	3
4.1	Instalando PyEnv no HPCC Marvin	3
4.2	Instalando as Dependências do Python	4
4.3	Enviando Imagens para o HPCC Marvin	4
4.4	Executando o Pipeline no HPCC Marvin	4
4.4.1	Monitoramento do Job SLURM	4
5	Visualizando os Resultados	5
6	Ajustando o pipeline parasitecolocalization.cppipe	5
6.1	Entendendo e alterando o arquivo cppipe para ajustes na análise	5
6.2	Acessando o VNC para usar o CellProfiler	6
6.2.1	Editando o arquivo cppipe no CellProfiler	6
6.2.2	Alterações comuns no arquivo cppipe	6
6.2.3	Testando o pipeline após alterações	7

1 Introdução

Este guia do usuário fornece instruções detalhadas sobre como executar a análise de viabilidade celular. O objetivo deste guia é auxiliar os usuários a configurar e executar o pipeline de análise, bem como a interpretar os resultados obtidos.

2 Requisitos

Para executar a análise de viabilidade celular, você precisará dos seguintes requisitos:

- **Acesso ao GitHub do CNPEM:** Você deve ter acesso ao repositório do projeto (<https://github.com/cnpem/CellViability>) no GitHub do CNPEM (<https://github.com/cnpem/>) para baixar os arquivos necessários. Se não tiver acesso, entre em contato com a Equipe de Dados Biológicos (edb@lnbio.cnpem.br).
- **Acesso ao HPCC Marvin:** Você deve ter acesso ao HPCC Marvin para executar o pipeline de análise. Se você não tiver acesso, acesse <https://marvindocs.cnpem.br/primeiros-passos/index.html>.

3 Acessando o HPCC Marvin

Para acessar o HPCC Marvin, comece abrindo o terminal. Se estiver usando Windows, abra o PowerShell; se estiver usando Linux ou MacOS, abra o Terminal. Para acessar o HPCC Marvin, use o comando:

```
ssh <seu.login.cnpem>@marvin.cnpem.br
```

Após inserir o comando, será solicitado que você insira sua senha, digite sua senha institucional.

4 Preparando o Repositório ParasiteCoLocalization

Para clonar o repositório [CellViability](#), execute o seguinte comando no terminal:

```
git clone https://github.com/cnpem/CellViability.git
```

O comando acima cria uma pasta chamada `CellViability` no diretório atual. Esta pasta contém todos os arquivos necessários para a execução da análise de viabilidade celular, incluindo o script `run.sh`, que automatiza a execução da análise via SLURM.

Após clonar o repositório `CellViability`, acesse o diretório do projeto com o comando:

```
cd CellViability
```

4.1 Instalando PyEnv no HPCC Marvin

O PyEnv é uma ferramenta que permite instalar e gerenciar várias versões do Python em um ambiente virtual. Para instalar o PyEnv no HPCC Marvin, execute os seguintes comandos no terminal:

```
curl https://pyenv.run | bash
echo 'export PATH="$HOME/.pyenv/bin:$PATH"' >> ~/.bashrc
echo 'eval "$(pyenv init -)"' >> ~/.bashrc
echo 'eval "$(pyenv virtualenv-init -)"' >> ~/.bashrc
source ~/.bashrc
```

Para instalar uma versão específica do Python, execute o seguinte comando:

```
pyenv install 3.8.19
```

Para definir a versão do Python como padrão neste diretório, execute o seguinte comando:

```
pyenv local 3.8.19
```

Caso deseje utilizar outra versão do Python, você pode selecionar a versão desejada para a sessão atual com o comando:

```
pyenv shell <versão_desejada>
```

4.2 Instalando as Dependências do Python

Para instalar as dependências do Python, execute o seguinte comando no terminal:

```
pip install -r requirements.txt
```

4.3 Enviando Imagens para o HPCC Marvin

Para enviar as imagens para o HPCC Marvin, você pode usar o comando `scp` em qualquer terminal.

```
scp /caminho/para/suas/imagens/*.tiff  
↪ <seu.login.cnpem>@marvin.cnpem.br:/caminho/para/CellViability/data
```

Caso você esteja usando o Windows, você pode usar o <https://winscp.net/eng/download.php> para transferir arquivos de e para o HPCC Marvin. Para isso, siga as instruções abaixo:

1. Abra o WinSCP e insira o endereço do servidor (`marvin.cnpem.br`), seu nome de usuário e senha.
2. Navegue até o diretório `CellViability/data` no HPCC Marvin.
3. Arraste e solte os arquivos de imagem do seu computador para o diretório `CellViability/data` no HPCC Marvin.

4.4 Executando o Pipeline no HPCC Marvin

Após enviar as imagens para o HPCC Marvin, você pode executar o pipeline de análise de viabilidade celular. O script `run.sh` executa o pipeline de análise via SLURM, que é o sistema de gerenciamento de tarefas usado no HPCC Marvin. Para isso, execute o script `run.sh` no diretório `CellViability`.

```
sbatch run.sh -m marvin
```

O comando acima retorna o ID do job SLURM, que você pode usar para monitorar o progresso da análise.

4.4.1 Monitoramento do Job SLURM

Para monitorar o progresso do job SLURM, você pode usar o comando `squeue` no terminal. O comando `squeue` exibe informações sobre os jobs SLURM em execução no HPCC Marvin.

```
squeue -u <seu.login.cnpem>
```

Se o job estiver em execução, o estado será exibido como R (em execução). Se o job estiver aguardando na fila, o estado será PD (pendente).

Caso deseje cancelar o job SLURM, você pode usar o comando `sancel` no terminal. O comando `sancel` cancela um job SLURM em execução no HPCC Marvin.

```
sancel <job_id>
```

Por fim, você pode verificar o arquivo de log gerado pelo SLURM (`slurm_<job_id>.out`) para obter informações sobre o seu job.

5 Visualizando os Resultados

Após a conclusão da análise, os resultados serão armazenados no diretório `results/`, que incluem os seguintes arquivos:

- `summary.csv`: Este arquivo contém um resumo dos resultados da análise por poço, incluindo contagem total de células.

	TotalCells
---	-----
A1	487.0
A2	217.0
A3	863.0
A4	58.0
A5	180.0
A6	54.0
A7	296.0
A8	227.0

- `plate_map/`: Este diretório contém os mapas interativos da placa de 384 poços (Figura 1), contendo:
 - `number_of_cells.html`: contagem total de células por poço.

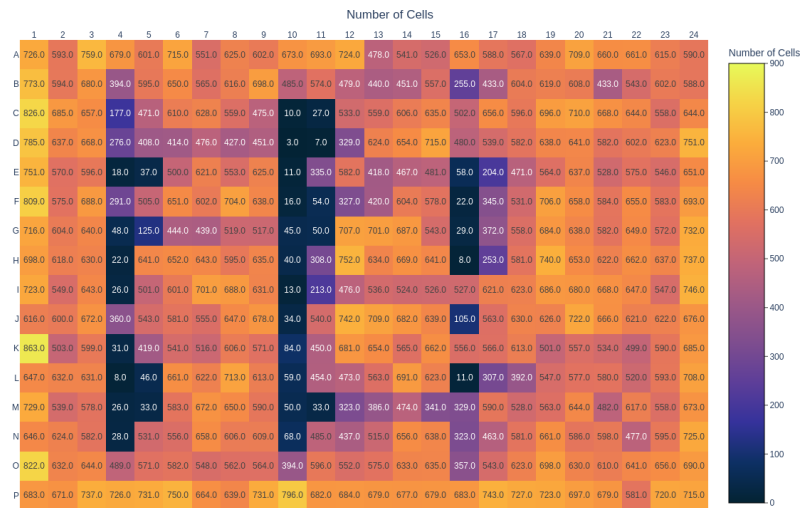


Figure 1: Exemplo de mapa interativo da placa de 384 poços.

6 Ajustando o pipeline `parasitecolocalization.cppipe`

6.1 Entendendo e alterando o arquivo `cppipe` para ajustes na análise

Dependendo dos dados que você está realizando, pode ser necessário ajustar o arquivo `cppipe` para configurar corretamente os módulos e as etapas do pipeline. Alterações no arquivo `cppipe` são importantes quando é necessário adaptar o pipeline para diferentes tipos de imagens ou análises.

Por exemplo, ao analisar células em imagens de alta resolução, pode ser necessário ajustar os parâmetros do módulo de segmentação para lidar com a maior resolução. Se estiver trabalhando com imagens de fluorescência, os parâmetros podem ser modificados para considerar o contraste ou a intensidade das células.

Para isso é preciso alterar o arquivo `cppipe`, que é um arquivo de configuração do CellProfiler que define a sequência de módulos a serem executados durante a análise de imagens, bem como os parâmetros específicos de cada módulo. Ele contém uma série de módulos que são executados sequencialmente, cada um realizando uma tarefa específica, como segmentação, medição ou visualização das imagens. Cada módulo possui parâmetros ajustáveis que controlam o comportamento da análise.

6.2 Acessando o VNC para usar o CellProfiler

Para acessar o VNC e configurar o ambiente para executar o CellProfiler, siga os passos abaixo:

1. Acesse o endereço <https://marvin.cnpem.br>.
2. Na aba **App Interativos**, selecione a opção **VNC**.
3. Na nova janela, configure os parâmetros:
 - Escolha o número de horas que pretende utilizar.
 - Configure o número de GPUs necessárias (**Exemplo: 1, 2 e 3**).
 - Selecione a opção **gui-gpu-small** como o perfil de GPU.
4. Clique em **Iniciar Sessão** para abrir o ambiente de VNC.

Após o ambiente ser inicializado, abra o terminal e execute o seguinte comando para rodar o CellProfiler:

```
singularity run --nv /opt/images/cellprofiler/cellprofiler-4_2_6.sif ~/Desktop/session.out
```

6.2.1 Editando o arquivo `cppipe` no CellProfiler

O CellProfiler oferece tanto uma interface gráfica quanto uma de linha de comando para editar arquivos `cppipe`. Abaixo, detalhamos os passos para realizar edições utilizando a interface gráfica:

1. **Abrir o CellProfiler:** Inicie o CellProfiler em sua máquina local ou no ambiente de trabalho.
2. **Carregar o pipeline existente:** No menu inicial do CellProfiler, clique em **Open Pipeline** e selecione o arquivo `cppipe` que deseja editar.
3. **Modificar módulos e parâmetros:** O CellProfiler exibirá todos os módulos do pipeline na interface. Clique em qualquer módulo para visualizar e modificar seus parâmetros.
 - Ajuste valores como **Thresholding** ou **Rescaling** no módulo de segmentação caso necessário.
 - Para adicionar novos módulos, clique em **Add Module** e escolha o módulo apropriado.
4. **Salvar o novo pipeline:** Após realizar as alterações, clique em **Save Pipeline As** e salve o arquivo modificado.
5. **Executar o pipeline:** Após salvar, execute o pipeline no CellProfiler ou via linha de comando, conforme descrito anteriormente.

6.2.2 Alterações comuns no arquivo `cppipe`

Dependendo do tipo de análise e das características das imagens, você pode realizar alterações como:

- Alterar o método de segmentação para lidar com diferentes características dos objetos nas imagens.
- Adicionar ou remover módulos para medir características específicas, como intensidade média das células.
- Ajustar os parâmetros de tamanho de objeto para considerar variações de tamanho das células.
- Modificar parâmetros de medição para analisar aspectos específicos, como área, intensidade ou forma.

6.2.3 Testando o pipeline após alterações

Após alterar o arquivo `cppipe`, é fundamental testá-lo com um conjunto representativo de imagens. Execute o pipeline em um subconjunto de imagens para garantir que os módulos estejam funcionando corretamente e que as medições estejam adequadas. Isso ajuda a identificar e corrigir rapidamente problemas antes de aplicar o pipeline ao conjunto completo de dados.