

**Deteção, reconhecimento e alerta de objetos em vídeo**

Licenciatura em Engenharia Informática

Marco António da Silva Ferreira

Leiria, « mês » de 2024



**Deteção, reconhecimento e alerta de objetos em vídeo**

Licenciatura em Engenharia Informática

Válter Santos Pedrosa

Trabalho de Projeto da unidade curricular de Projeto Informático realizado sob a orientação do Professor Doutor Paulo Costa e do Professor Doutor Filipe Neves

Leiria, « mês » de 2024

# Dedicatória

Inserir aqui a dedicatória. Trata-se de um elemento **facultativo**.

Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória.

Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória. Texto da dedicatória.

# Agradecimentos

Inserir aqui os agradecimentos. Trata-se de um elemento **facultativo**.

Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos.

Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos. Texto dos agradecimentos.

# Resumo

Inserir aqui o resumo. Trata-se de um elemento **obrigatório**.

Deve começar sempre numa página ímpar. Se ocupar um número par de páginas (p. ex. 2), deve ajustar-se o texto para que a próxima secção (abstract) se inicie numa página ímpar. O resumo deve acabar com a lista de palavras-chave.

**No resumo deve dar-se nota das principais ideias do trabalho (objetivos e conclusões).**

Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo.

Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo. Texto do resumo.

**Palavras-chave:** deteção, yolo, reconhecimento, IA.

# Abstract

Please insert here the abstract in English. This is a **mandatory** element.

The abstract should always start in an odd page. If the length is a multiple of two, the text should be adjusted in order to the next section start also in an odd page. The abstract should end with a list of keywords.

Please insert here the abstract in English. Please insert here the abstract in English. Please insert here the abstract in English. Please insert here the abstract in English. Please insert here the abstract in English.

Please insert here the abstract in English. Please insert here the abstract in English. Please insert here the abstract in English. Please insert here the abstract in English. Please insert here the abstract in English.

**Keywords:** maximum of 6 words separated by “,”

Índice

Trata-se de um elemento **obrigatório**. Nota: **o índice nunca figura do índice.**

[Dedicatória ii](#_Toc163578258)

[Agradecimentos iii](#_Toc163578259)

[Resumo iv](#_Toc163578260)

[Abstract v](#_Toc163578261)

[Lista de Figuras ix](#_Toc163578262)

[Lista de tabelas x](#_Toc163578263)

[Lista de siglas e acrónimos xi](#_Toc163578264)

[1. Introdução 1](#_Toc163578265)

[2. Título do capítulo 2](#_Toc163578266)

[2.1. Título da secção 4](#_Toc163578267)

[2.2. Título da secção 4](#_Toc163578268)

[2.2.1. Título da subsecção 5](#_Toc163578269)

[2.2.2. Título da subsecção 5](#_Toc163578270)

[2.2.3. Título da subsecção 5](#_Toc163578271)

[2.3. Título da secção 5](#_Toc163578272)

[3. Testes 6](#_Toc163578273)

[3.1. Teste de margem ideal para definir posição do objeto 6](#_Toc163578274)

[3.1.1. Introdução 6](#_Toc163578275)

[3.1.2. Condições do teste 6](#_Toc163578276)

[3.1.3. Resultados obtidos: 7](#_Toc163578277)

[3.1.3.1. Class “backpack” 7](#_Toc163578278)

[3.1.3.2. Class “sports ball” 8](#_Toc163578279)

[9](#_Toc163578280)

[9](#_Toc163578281)

[3.1.3.3. Class “Person” 9](#_Toc163578282)

[10](#_Toc163578283)

[3.1.4. Conclusão: 13](#_Toc163578284)

[4. Conclusões ou Conclusão 14](#_Toc163578285)

[Bibliografia ou Referências Bibliográficas 15](#_Toc163578286)

[Anexos 16](#_Toc163578287)

[Glossário 17](#_Toc163578288)

# Lista de Figuras

Elemento a figurar, **quando aplicável**.

[Figura 2.1 - Texto ilustrativo da figura 1. 3](#_Toc164784383)

[Figura 2.4 - Texto ilustrativo da figura 2. 3](#_Toc164784384)

[Figura 3.1 - Gráfico cantos backpack a 5M 7](#_Toc164784385)

[Figura 3.2 - Gráfico cantos backpack a 10M 8](#_Toc164784386)

[Figura 3.3 - Gráfico cantos backpack a 15M 8](#_Toc164784387)

[Figura 3.4 - Gráfico cantos bola a 5M 9](#_Toc164784388)

[Figura 3.5 - Gráfico cantos bola a 10M 9](#_Toc164784389)

[Figura 3.6 - Gráfico cantos bola a 12M 10](#_Toc164784390)

[Figura 3.7 - Gráfico cantos person a 5M 10](#_Toc164784391)

[Figura 3.8 - Gráfico cantos person a 35M 11](#_Toc164784392)

[Figura 3.9 - Gráfico cantos person a 70M 11](#_Toc164784393)

[Figura 3.10 - Gráfico performance parallel 13](#_Toc164784394)

[Figura 3.11 - Gráfico performance parallel 5 segundos 14](#_Toc164784395)

[Figura 3.12 - Gráfico performance parallel 10 segundos 15](#_Toc164784396)

[Figura 3.13 - Gráfico performance open & close 15](#_Toc164784397)

[Figura 3.14 - Gráfico performance open & close threading 16](#_Toc164784398)

# Lista de tabelas

Elemento a figurar, **quando aplicável**.

[Tabela 1.1 - Texto ilustrativo da tabela 1. 3](#_Toc92389035)

# Lista de siglas e acrónimos

Elemento a figurar, **quando aplicável**.

|  |  |
| --- | --- |
| ESTG | Escola Superior de Tecnologia e Gestão |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Cuidados na elaboração da lista de siglas e acrónimos:

* Ordenação alfabética;
* Apenas as que sejam relevantes para a leitura do texto.

Adicionar mais entradas à tabela, caso seja necessário (a tabela não tem contornos, mas está no texto).

# Introdução

Introdução deve conter resumidamente os seguintes elementos:

* O objeto do trabalho (o tema);
* A justificação ou a pertinência do tema;
* Os objetivos do trabalho (gerais e específicos, perguntas a responder ou hipóteses a testar);
* Os métodos e as técnicas utilizados;
* Estrutura do trabalho.

# Título do capítulo

Aqui, inicia-se o desenvolvimento dos trabalhos e deve indicar-se o título do capítulo primeiro.

O desenvolvimento do trabalho deve ser adequado à natureza da unidade curricular (dissertação/trabalho de projeto/relatório de estágio) e deve seguir as práticas mais disseminadas na área em causa.

Estrutura: pode ter, por exemplo, capítulos, secções e subsecções.

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

De seguida, apresenta-se um exemplo de como as imagens devem ser colocadas no texto:

* Centradas;
* Numeradas;
* Referenciadas no texto;
* Sem contornos;
* Eventualmente, colocar a(s) figura(s) numa tabela para melhorar a formatação;
* Usar sempre o estilo “caption” para o texto das legendas;
* Identificação da figura em “negrito” (o estilo trata desta parte). Texto da legenda não deve estar em negrito.

|  |
| --- |
| Figura 2.1 - Texto ilustrativo da figura 1. |

De seguida apresenta-se a figura 2.

|  |
| --- |
| Figura 2.4 - Texto ilustrativo da figura 2. |

Cuidados com a inserção de tabelas:

* Centradas;
* Numeradas;
* Referenciadas no texto;
* Usar sempre o estilo “caption” para o texto das legendas;
* Identificação da tabela em “negrito” (o estilo trata desta parte). Texto da legenda não deve estar em negrito.

Tabela 1.1 - Texto ilustrativo da tabela 1.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Continuação do texto do desenvolvimento.

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

## Título da secção

A figurar no caso de o capítulo ter várias secções. Nota: apenas devemos criar secções e subsecções quando existem mais do que uma.

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

## Título da secção

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

### Título da subsecção

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

### Título da subsecção

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

### Título da subsecção

Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento. Texto do desenvolvimento.

## Título da secção

# Testes

## Teste de margem ideal para definir posição do objeto

### Introdução

O YOLO, após detetar um objeto, desenha uma bounding box que delimita o objeto detetado. Mesmo que o input não se altere, no caso deste projeto em que o input são camaras de vídeo, significa objeto na mesma posição, esta box pode variar ligeiramente devido a vários fatores. Estes fatores compreendem: variabilidade de rede, anchor boxes ou Non-Maximum Supression (NMS). Variabilidade de rede porque, a rede YOLO, como qualquer rede neuronal de deep lerning, tem variabilidade inerente nas suas previsões, devido à variação de valores inicias de peso, processos de otimização e descendente estocástico de gradiente (Stochastic Gradient Descent - SGD).Entende-se por Anchor boxes as formas e tamanhos predefinidos pelo algoritmo usadas para prever as bouding boxes. Se as anchor boxes não corresponderem exatamente à forma e tamanho dos objetos na imagem, pode também levar a variações nas bounding boxes. Por fim, NMS é uma técnica de pós-processamento usada pelo YOLO para filtrar as bounding boxes previstas. Neste processo é escolhida a box com maior grau de confiança e suprime-se as boxes sobrepostas com menor grau de confiança. O limite, isto é, o valor de confiança mínimo necessário para que uma bounding box seja considerada válida, usado nesta técnica pode afetar que bounding boxes são selecionadas, resulta em variações nas previsões finais.

Posto isto, submetemo-nos à criação de um cenário de testes para definir qual a margem ideal a definir para compensar a variação da posição da bounding box em objetos parados. Deste modo garantimos uma deteção de movimento mais rigorosa, e consequentemente alertas de segurança mais precisos e confiáveis.

### Condições do teste

Para avaliar qual a margem que confere a melhor exatidão dos resultados, foram simulados 17 cenários de testes variando os objetos a detetar e a distância a que se encontram do dispositivo de gravação. A posição da camara manteve-se a mesma durante todo o procedimento bem como o cenário onde se encontraram os objetos a detetar. Este cenário consistia num dia de sol, céu limpo com presença de sombra onde os objetos foram detetados.

Os cenários de teste foram os seguintes:

* Mochila (backpack) a 5,10 e 15 metros
* Bola (sports ball) a 5,10 e 12 metros
* Pessoa (person) a 5,10,15,20,25,30,35,40,50,60 e 70 metros

### Resultados obtidos:

Os gráficos seguintes representam, à esquerda, a distância entre os cantos inferior esquerdo (canto 1) do frame anterior e o atual e a distância do canto superior direito (canto 2) entre o frame anterior e atual à direita.

### Class “backpack”

Uma imagem com texto, file, diagrama, Gráfico

Descrição gerada automaticamente

Figura 3.1 - Gráfico cantos backpack a 5M

Uma imagem com texto, file, diagrama, Gráfico

Descrição gerada automaticamente

Figura 3.2 - Gráfico cantos backpack a 10M

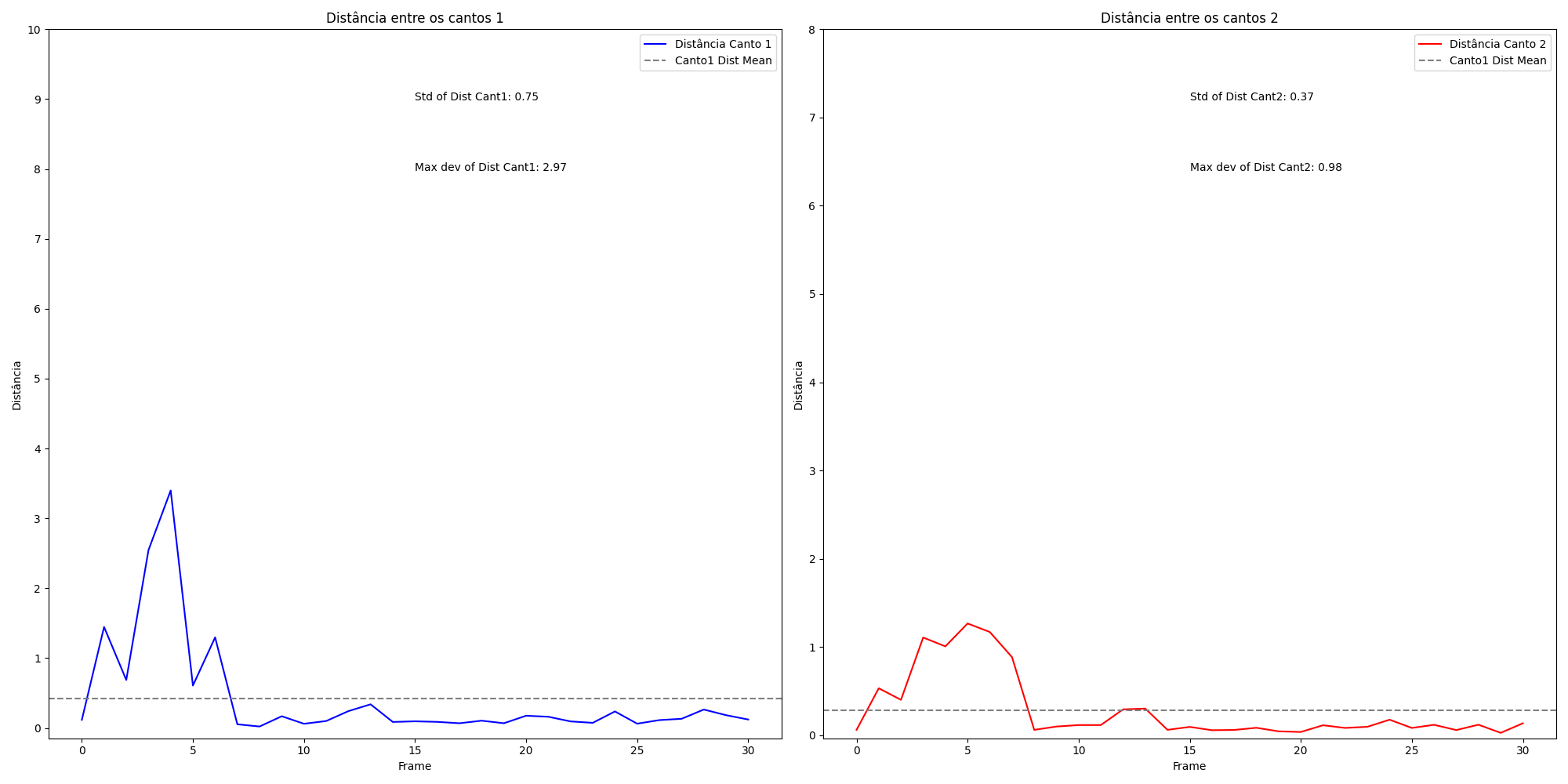


Figura 3.3 - Gráfico cantos backpack a 15M

### Class “sports ball”

Uma imagem com diagrama, file, texto, Gráfico

Descrição gerada automaticamente

Figura 3.4 - Gráfico cantos bola a 5M

Uma imagem com texto, file, Gráfico, diagrama

Descrição gerada automaticamente

Figura 3.5 - Gráfico cantos bola a 10M

Uma imagem com texto, file, diagrama, Gráfico

Descrição gerada automaticamente

Figura 3.6 - Gráfico cantos bola a 12M

### Class “Person”

Uma imagem com file, texto, diagrama, Gráfico

Descrição gerada automaticamente

Figura 3.7 - Gráfico cantos person a 5M

Uma imagem com file, diagrama, texto, Gráfico

Descrição gerada automaticamente

Figura 3.8 - Gráfico cantos person a 35M

Uma imagem com texto, file, diagrama, Gráfico

Descrição gerada automaticamente

Figura 3.9 - Gráfico cantos person a 70M

### Conclusão:

Após teste dos vários cenários, seguidos da análise e discussão dos resultados obtidos, consideramos 4 px como margem ideal a aplicar para desprezar as possíveis variações das bounding boxes geradas pelo modelo de deteção usado em objetos parados. Isto deve-se ao facto dos maiores desvios obtidos serem na ordem dos 3px, sendo o máximo deles 3.15 no caso de teste backpack a 5 metros. Sendo 4 o próximo inteiro, 4px cobre todos os desvios que obtivemos na fase de testes.

## Teste da arquitetura para recolher frames de dispositivos

### Introdução

Serve o presente teste para testar qual a arquitetura que melhor se adequa à nossa aplicação e aos seus requisitos. Para isso temos 4 modelos para teste sendo eles:

### Grab & Retrieve

Neste modelo, após inicializarmos as câmaras com o videCapture de cada dispositivo, criamos uma thread de captura por dispositivo. Esta thread apenas agarra (grab) continuamente frames do dispositivo para posterior descodificação e análise do frame. Na thread principal, de X em X segundos (definido pelo utilizador) chamamos a função retrieveFrames que percorre o vetor das câmaras previamente inicializadas e recolhe (retrieve) o último frame agarrado (grabbed) pela thread de captura do dispositivo associado descodificando-o. De seguida enviamos cada frame para a função YOLO para o processar.

### Parallel

Nesta arquitetura, para cada dispositivo de captura selecionado, é criada uma thread que lê, descodifica e envia o frame ao YOLO para detetar objetos na imagem.

### Open & Close

Neste padrão (ausente de threads) cada dispositivo de captura é inicializado, captura o frame, processa-o para deteção através do YOLO e fecha o dispositivo, seguindo para o próximo continuamente, N vezes, sendo N o número de dispositivos selecionado.

### Open & Close threading

Neste modelo, inicializa-se os dispositivos e cria-se uma thread por dispositivo para ler um frame. Após todas as threads terem lido um frame, cria-se uma thread por frame para o processar no YOLO. Finalmente, depois das threads terminarem o processamento, fecham-se os dispositivos inicializados no início.

### Condições de teste

Para comparar os vários modelos de implementação, utilizamos a biblioteca “psutil” do python que nos permite avaliar a % de CPU usage bem como a % de RAM usage durante todo o processo. Para avaliar os cenários, utilizamos as mesmas 5 câmaras, das quais 2 locais e 3 por IP e definimos 40 iterações para avaliar o tempo de execução.

### Resultados obtidos

### Grab & Retrieve

Neste padrão não foi possível concluir os testes porque ao usar câmaras IP em vez de locais no videoCapture, o retrieve não consegue descodificar (retrieve) o frame agarrado (grabbed) anteriormente, pelo que excluímos das opções.

### Parallel

Uma imagem com texto, file, diagrama, captura de ecrã

Descrição gerada automaticamente

Figura 3.10 - Gráfico performance parallel

Ao analisar a percentagem elevada de CPU usage nesta arquitetura, acrescentamos dois novos casos de teste para tentar diminuir a % de CPU usage sacrificando tempo de execução. Para isso, adicionamos um time.sleep de 5 e 10 segundos.

### Uma imagem com texto, file, captura de ecrã, diagrama Descrição gerada automaticamenteParallel 5 sec

Figura 3.11 - Gráfico performance parallel 5 segundos

### Uma imagem com texto, file, diagrama, Gráfico Descrição gerada automaticamenteParallel 10 sec

Figura 3.12 - Gráfico performance parallel 10 segundos

### Uma imagem com texto, file, Gráfico, diagrama Descrição gerada automaticamenteOpen & Close

Figura 3.13 - Gráfico performance open & close

### Open & Close threading

Uma imagem com texto, file, diagrama, Gráfico

Descrição gerada automaticamente

Figura 3.14 - Gráfico performance open & close threading

### Discussão dos Resultados

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Método | Tempo execução | Máx – Init CPU % | Máx – Init RAM % | Mean – Init CPU % | Mean – Init RAM % |
| Parallel | 1m 44s | 100 | 10,80 | 85,88 | 10,03 |
| Parallel 5 sec | 5m 24s | 100 | 8,50 | 34,55 | 7,67 |
| Parallel 10 sec | 7m 39s | 35,20 | 8,30 | 20,46 | 4,11 |
| Open & Close | 17m 58s | 28,30 | 19,50 | 9,86 | 13,53 |
| Open & Close threading | 10m 25s | 32,40 | 13 | 18,14 | 8,77 |

Através dos dados retirados podemos criar a seguinte tabela:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Menor tempo execução | Maior tempo  Execução | Menor CPU usage | Maior CPU usage | Menor RAM usage | Maior RAM usage |
| Parallel | Open & Close | Open & Close | Parallel | Parallel 10 sec | Open & Close threading |

Neste projeto, priorizamos o menor custo computacional em relação ao tempo de execução. Dito esto, excluímos o modelo Parallel pois utiliza o esforço máximo do CPU e o Open & Close threading por representar o maior esforço sobre a RAM.

Entre o Parallel 5 sec e o Parallel 10 sec, o último representa o melhor comparativamente porque demora 1,42 vezes mais para concluir o teste mas utiliza menos 1,69 vezes o CPU e 1,87 vezes menos a RAM.

O Open & Close é o que apresenta menor custo para o processador (9,86%) mas demora demasiado tempo em relação ao Parallel 10 sec que utiliza 20,46% do processador, o que representa 2,08 vezes mais que o anterior, mas acaba o teste 2,35 vezes mais cedo, o que representa um rácio maior de tempo / CPU usage, pelo que também exclui o Open & Close em comparação ao Parallel 10 sec.

### Conclusão

Após análise e discussão dos resultados obtidos, escolhemos a arquitetura Parallel 10 sec como a mais adequada para a nossa aplicação por ser o mais equilibrado quando aplicamos os nossos critérios.

# Conclusões ou Conclusão

Inserir aqui as conclusões ou conclusão. Trata-se de um elemento **obrigatório**.

A conclusão:

* Deve ser sucinta;
* Não deve conter informações ou ideias novas;
* Deve permitir concluir se se atingiram os objetivos enunciados na introdução.

Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão.

Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão. Texto da conclusão.

# Bibliografia ou Referências Bibliográficas

Inserir aqui a bibliografia ou referências bibliográficas. Trata-se de um elemento **obrigatório**.

Notas: o sistema a adotar para a apresentação das referências bibliográficas e as suas citações deve:

* Respeitar uma norma estabelecida;
* Seguir as práticas mais disseminadas na área em causa;
* Ser empregue de modo uniforme em todo o documento.

Bibliografia – quando se coloca toda a bibliografia consultada;

Referências bibliográficas – quando se faz referência apenas à bibliografia citada.

# Anexos

Elemento a figurar, **quando aplicável**.

# Glossário

Elemento a figurar, **quando aplicável**.