

Projeto de Processamento de Imagem e Classificação de Veículos - Carvana

Anabela Reigoto, *Estudante, FEUP*, André Oliveira, *Estudante, FEUP* e Baltasar Aroso, *Estudante, FEUP*

Sumário - Este projeto consistiu no desenvolvimento e teste da manipulação de imagens fornecidas por um stand de automóveis, Carvana, através de métodos de pré-processamento de imagem, segmentação, análise e reconhecimento e classificação para o projecto final da unidade curricular Sistemas Baseados em Visão.

Palavras-Chave – Segmentação, Classificação, Reconhecimento.

I. INTRODUÇÃO

Este relatório tem como objetivo complementar o código efetuado para o projeto final da Unidade Curricular Sistemas Baseados em Visão, elucidando a sua natureza mais teórica, visto que esta poderá não ter sido bem detalhada na apresentação.

Este trabalho tinha como primeiro objetivo a implementação de um algoritmo em MATLAB capaz de segmentar as imagens apresentadas para que se assemelhassem às máscaras *ground-truth* fornecidas, sendo o seu termo de comparação o cálculo da eficiência através do coeficiente de *Sørensen–Dice*.

Numa segunda fase, era pedida a classificação dos veículos nas suas possíveis classes: Compacto, Pick-up, Sedan ou SUV; sendo possível recorrer, para este efeito, às suas máscaras *ground-truth*. Para a realização desta tarefa havia livre-arbítrio para escolher qual a melhor disposição do carro para que o algoritmo fosse mais eficiente. Posto isto, após uma análise das imagens fornecidas escolhemos a vista lateral.

A última tarefa tinha como objetivo a classificação dos veículos na sua marca correspondente, dentro das 7 marcas possíveis: Mini, BMW, Chevrolet, Ford, Honda, Jaguar e Audi. Para a realização deste algoritmo escolhemos a disposição traseira dos automóveis, com a exceção de um Mini *Clubman* cujo símbolo da marca se encontrava no canto inferior esquerdo da traseira, o que impossibilitava a sistematização do código que procurava o símbolo na parte central da imagem, tendo sido assumido *a priori* que pelo menos uma das deteções estaria errada.

II. SEGMENTAÇÃO

Numa primeira abordagem, tentou-se remover o letreiro CARVANA recorrendo a uma operação de **close** com um

elemento estruturante “linha”, horizontal, grande o suficiente para que estas desaparecessem completamente (150), tentando-se de seguida proceder à *edge detection* (deteção de orla) com a função **edge**, usando o algoritmo de *Canny* e um *threshold* adequado. Os resultados ficaram aquém dos esperados, visto que, após a remoção das letras, o fundo apresentava bastantes mudanças súbitas na intensidade da cor, semelhante à imagem ter sido esboçada perfeitamente na horizontal, devido ao elemento estruturante escolhido.

Numa segunda abordagem, e percebendo desta vez que a imagem requeria algum pré-processamento mais elaborado antes de se tentar a deteção de orla, decidiu-se aplicar a operação de **close**, desta vez com uma linha ligeiramente mais curta (120), de modo a remover em grande parte as letras, mas não danificando em demasia as orlas do carro. De seguida, após aplicar um filtro gaussiano para remoção de algum ruído, efetuou-se a soma de todas as intensidades, em cada coluna da matriz obtida da leitura da imagem.¹ Isto teve como objetivo a análise de máximos e mínimos para que se pudesse tentar encontrar as coordenadas limite do veículo, seja esse qual for. Este procedimento procura “encontrar” um retângulo que limite o veículo, tornando tudo à volta deste completamente uniforme (escolheu-se a cor cinza).

Após uma primeira observação de máximos e mínimos, encontrou-se um problema: mesmo que se recorresse a uma remoção quase completa das letras (elemento estruturante ‘linha’ de comprimento maior), as transições abruptas do fundo branco para as letras escuras permaneciam. A solução que se provou bastante eficaz consistiu numa medição das larguras de letras relevantes (‘C’, ‘R’ e ‘N’: letras que possuem fortes traços verticais, ou quase verticais), ainda na análise de máximos e mínimos. Verificou-se que, mesmo para fotografias com diferente distância à câmara fotográfica, essas larguras se encontravam em 2 intervalos restritos² e facilmente detetáveis, uma vez que correspondiam sempre a um máximo precedido de um mínimo (claro; escuro; claro). A sua remoção torna-se, portanto, trivial: procuram-se os máximos e mínimos (guardados separadamente) e, sempre que houver um máximo com um mínimo precedido a uma distância aproximadamente igual à largura de uma letra, ambos são removidos. São ainda removidos todos os máximos e mínimos abaixo de um certo nível, relativo ao maior máximo. Neste momento, o algoritmo encontra-se pronto para determinar o limite esquerdo e o limite

¹ Todas as imagens lidas pelo algoritmo em MATLAB foram, aquando da sua leitura, convertidas de RGB para escala de cinza (rgb2gray) e para dupla precisão (im2double).

² Presentes no código MATLAB nas variáveis *lw_sup* e *lw_inf*, que contêm, respetivamente, os limites superiores e inferiores desses intervalos. Os valores que serviram de guia para esses intervalos encontram-se acima dessas variáveis, comentados.

direito do veículo através do respetivo pico (máximo ou mínimo) extremo. Para encontrar os limites superior e inferior usou-se um procedimento semelhante, sendo que para estes foi possível desprezar o impacto das letras fazendo a operação de **close** com o elemento estruturante ‘linha’, comprimento 150, e aplicando um filtro gaussiano mais potente.

Tendo agora as dimensões da caixa que delimita o carro, aproximadamente, e aplicando-lhe uma margem de segurança obtida através de experimentação, colocou-se o fundo exterior a este retângulo todo a cinza. Aplica-se um algoritmo de detecção de orla de *Canny* e remove-se a orla detetada em volta do retângulo (visto que é indesejada e facilmente removível). Através da transformada de *Hough*, podemos detetar segmentos de reta na imagem correspondentes à orla das letras e a traços verticais irrelevantes, tendo em conta a sua inclinação e definindo um tamanho bom o suficiente para não apanhar traços importantes do carro, mas que apanhe orlas de letra, dentro de um limite inferior razoável. Utilizou-se o tamanho 50 e os ângulos descritíveis nos vetores [-40:-34 34:40 -26:-18 18:26] (diagonais) e [-3:3] (verticais).³

No fim, com vista a “limpar” e preencher o segmento do veículo obtido, aplicou-se uma operação binária de **thin**, seguida de um **close**, depois **fill**, e por fim um **open**.

O procedimento acima descrito encontra-se dentro do script **task1.m**, na secção *Segmentation*. Esta secção efetua a segmentação de todos os veículos em modo automático, sendo que no final nos indica a eficiência média obtida através da média dos coeficientes de *Sørensen–Dice* individuais. O script faz uso de 3 funções: **startup.m** (recebe as imagens de todos os carros em escala de cinza), **getGroundTruths.m** (recebe todos os ground-truths) e **bresenham.m**. Esta última foi obtida através da biblioteca pública de funções MATLAB e retorna todos os pontos de um segmento de reta limitado pelos pontos passados como argumentos, usando uma versão do algoritmo de *Bresenham* otimizada para MATLAB (sem *loops*).

Para o cálculo dos coeficientes de *Sørensen–Dice* individuais usou-se:

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (1)$$

Em (1), *TP* são os verdadeiros positivos, *FP* os falsos positivos e *FN* os falsos negativos.

Na Tabela 1 encontram-se os valores obtidos para a eficiência e gráficos que permitem uma boa análise quantitativa e detalhada da mesma. A eficiência média obtida foi de 94.55%.

III. CLASSIFICAÇÃO DOS VEÍCULOS NAS SUAS CLASSES

Os veículos da base de dados apresentada encontram-se separados em quatro classes: compactos, pick-ups, sedans e SUVs. A classificação dos veículos nas suas classes rege-se essencialmente pelas suas características dimensionais, sendo do senso comum normal de analisar que um veículo das classes pick-up e SUV é maior do que um veículo das outras duas classes. Porém, essas classificações são inconsistentes e

ambíguas. Pela análise das imagens da base de dados podemos verificar que a distância dos veículos à câmara de onde foi tirada a fotografia é variável. Posto isto e para garantir uma correta classificação dos dados, fez-se uma análise pormenorizada de todas as características que, na nossa opinião, seriam preponderantes para a completa caracterização dos veículos.

Numa primeira fase, os parâmetros analisados foram: a altura, a largura, o perímetro dado pelo retângulo que delimitava a altura e a largura do veículo e a área desse mesmo retângulo (box⁴). Foi calculado o valor máximo, mínimo e médio de todos estes parâmetros, como é apresentado nas Tabelas 2, 3, 4 e 5.

Perante a análise visual das imagens e dos resultados tabelados foi escolhida apenas a vista lateral para a classificação dos veículos.

Primeiramente procurou-se distinguir os veículos com base nas características mais intrínsecas à sua classe, como por exemplo a distinção das Pick-up pela altura da sua suspensão ou a separação dos SUVs pela extensão da sua “cúpula” face à sua largura. Para isto foram usadas as funções **suspension.m** e **decision_suspension.m** para a distinção das pick-ups e a função **decision_dome.m** para a distinção dos SUVs.

A função **suspension.m** calcula o tamanho da suspensão mínima e máxima de cada classe considerando a distância entre o último pixel branco no eixo central vertical da imagem e a altura delimitada pelos pneus do automóvel. Após a análise dos resultados apresentados na Tabela 6 verificou-se que o tamanho mínimo da suspensão das Pick-up era superior a todos os outros e que havia uma grande discrepância entre o tamanho da suspensão das Pick-up e SUVs face aos Compactos e Sedans. Posto isto, o algoritmo já conseguia separar primeiro as Pick-up e depois as SUVs. Sendo o último parâmetro de desempate entre os Compactos e os Sedans determinado pela relação dos seus eixos vertical e horizontal, altura e largura, respetivamente. Portanto, pela divisão da altura pela sua largura observou-se, como seria de esperar dado o extenso comprimento dos Sedans, que o valor mínimo desta divisão na classe Compactos era superior ao valor máximo na classe Sedans, concluindo desta maneira a classificação de todos os veículos nas quatro classes.

Porém e devido ao facto dos valores utilizados para a separação do SUV usando o **decision_dome.m** terem sido inseridos com base em tentativas e a margem de distinção ser muito pequena, uma outra opção para separar as SUVs seria o cálculo da sua área, usando a função do MATLAB **regionprops**, e inserindo como argumento ‘Area’.

Após a utilização e exploração da função **regionprops** e tendo em conta que os métodos que estavam a ser usados eram muito restritos à precisão da segmentação, procurou-se alternativas. Para isso aumentou-se a análise dos dados como é apresentado na Tabela 7. Nesta tabela são apresentados 4 parâmetros: área, formula, axis e box.

A formula corresponde ao quociente do quadrado do perímetro com a área. Este parâmetro é usado logo numa fase

³ Os ângulos das letras foram medidos com recurso à ferramenta de visualização de imagem do MATLAB e através de conhecimento básico matemático, analisando pontos dessas mesmas letras.

⁴ Este será o nome utilizado para a área do retângulo calculada pelo produto da altura com a largura em contraste com a área posteriormente mencionada como a soma de todos os pixels brancos na imagem do ground-truth.

inicial da árvore de decisão em três dos quatro algoritmos, uma vez que possui uma grande margem de decisão, cerca de 25.80%, que distancia as Pick-up dos restantes veículos. Este resultado é facilmente explicado tendo em conta que as Pick-up apresentam um tamanho elevado e uma “caixa aberta” que lhes aumenta o perímetro comparativamente aos veículos mais arredondados. A simples relação entre a soma dos catetos de um triângulo retângulo e a sua hipotenusa demonstra bem essa situação.

$$formula = \frac{Perimeter^2}{Area} \quad (2)$$

O axis é calculado através do quociente entre a altura e a largura e permite identificar os carros mais baixos e compridos, como os Sedans, dos carros mais altos e pequenos, como os Compactos.

$$axis = \frac{height}{width} \quad (3)$$

A box é calculada pelo produto da altura com a largura, correspondendo à área do retângulo definido pela *BoundingBox*. Este parâmetro é usado apenas no primeiro algoritmo para a separação entre os Sedans e os SUVs, dado o seu porte face à pequena altura dos Sedans e uma largura semelhante.

$$box = height \times width \quad (4)$$

Depois de uma análise detalhada dos dados obtidos e procurando criar uma árvore de decisão robusta aos erros da segmentação foram calculadas as margens de decisão tendo em conta os seis parâmetros apresentados na Tabela 5.

Perante as combinações nessa tabela assinaladas criou-se quatro algoritmos diferentes para a determinação da classe dos veículos. Todos os algoritmos apresentaram uma eficiência de 100% usando as imagens *ground-truth* na vista lateral.

Na utilização do algoritmo para a classificação das imagens obtidas com a segmentação da tarefa 1 obtivemos as eficiências apresentadas na Tabela 6. A eficiência mais elevada foi de apenas 88% no Algoritmo 1, sendo que todos os algoritmos demonstraram uma má classificação na classe das Pick-up. Como se pode verificar na análise dos resultados das eficiências pelas classes da Tabela 1, as Pick-ups apresentam a eficiência mais baixa. Para além disso, na mesma tabela podemos verificar que os veículos da vista lateral são os que sofrem uma pior segmentação. Posto isto, é normal ocorrer uma diminuição da eficiência dado que os critérios de decisão são bruscamente alterados pela segmentação.

Primeiramente a intenção da criação de quatro algoritmos seria aumentar a robustez da classificação face às imperfeições da segmentação, porém, após uma análise dos resultados obtidos verificou-se que nenhuma combinação seria melhor do que o Algoritmo 1, uma vez que todas as combinações erravam nos mesmos aspetos, a separação dos Compactos com os Sedans, e a falha na classificação das Pick-up.

IV. CLASSIFICAÇÃO DOS VEÍCULOS NAS SUAS MARCAS

Os veículos da base de dados fornecida apresentam sete marcas diferentes: Mini, BMW, Chevrolet, Ford, Honda, Jaguar e Audi. A classificação dos veículos nas suas marcas baseia-se na comparação do símbolo presente na traseira do carro com os símbolos das diferentes marcas. Para tal foi criada uma base de dados com os símbolos das sete marcas.

Na análise das imagens, podemos verificar que na traseira do carro, o símbolo da marca encontra-se maioritariamente no centro, à exceção do Mini *Clubman*, tendo esta localização pequenas variações entre diferentes classes de veículos.

Primeiramente, utilizou-se o algoritmo desenvolvido na tarefa 2 para distinguir os veículos em classes, utilizando as imagens *ground-truth*. Posteriormente selecionou-se a janela de corte consoante a classe do veículo em causa. Para o pré-processamento da imagem foi usado o filtro Laplaciano de Gauss com um sigma de 0.4 para as imagens a classificar e de 0.5 para as imagens da base de dados criada para comparação. Com isto procurou-se delimitar as linhas que definem o logótipo das marcas para que a correlação fosse mais precisa.

Na análise da correlação observou-se que os valores oscilam muito de marca para marca, sendo esta a razão de uma eficiência de 62.5%. O facto de alguns logótipos, nomeadamente o do Chevrolet, apresentarem uma correlação média superior na maioria dos carros, influencia a decisão do máximo valor da correlação.

Para uma maior uniformização dos resultados da convolução tendo em conta o mesmo símbolo para diferentes carros, tentou-se melhorar a classificação ao calcular o máximo valor da correlação de cada símbolo para os 16 carros e dividir esse vetor por esse máximo, procurando normalizar os valores. Porém o resultado obtido não foi o pretendido e o valor da eficiência diminuiu para 43.75%. Depois da análise dos resultados percebeu-se que vetores de correlações com valores menos distantes pioram a classificação. Por exemplo, considerando dois vetores, um com um primeiro máximo muito maior que o segundo, que resulta numa normalização em que o segundo máximo é por exemplo 0.5; e outro vetor onde há um conjunto de máximos relativos próximos do máximo absoluto, cuja normalização se sobrepõe aos 0.5 apesar desse não ser a escolha acertada.

V. CONCLUSÃO

Em suma, o processo mais complexo do projeto baseou-se na segmentação das imagens, que exigiu um maior teste e pré-processamento dos veículos. A classificação é, neste caso, muito influenciada pela segmentação por ela antecedida, sendo que deve ter por base o seu estudo na criação da árvore de decisão. Sendo certo que a eficiência dos algoritmos apresentados seria muito maior caso estes mesmos fossem feitos tendo em conta a análise das imagens obtidas pela segmentação, visto que esta pode cortar características determinantes do veículo. A última tarefa remete para uma classificação de outro tipo, mas que é fundamentada nas duas tarefas anteriores para que haja um resultado mais eficaz. Esta tem por base os resultados de uma correlação que falhou pela

falta de pré-processamento e uma pouca consistência na disposição dos logotipos no centro das imagens.

APÊNDICE

Veículo		1	2	3	4	5	Total
Compact	Front	97,31%	95,25%	96,27%	96,91%	-	96,43%
	Back	96,08%	91,44%	96,14%	93,01%	-	94,17%
	Side	89,15%	93,68%	95,74%	96,23%	-	93,70%
	Other	95,37%	93,09%	90,89%	98,07%	-	94,36%
	Total	94,48%	93,37%	94,76%	96,06%	-	94,67%
Pick-up	Front	96,42%	97,58%	92,39%	-	-	95,46%
	Back	94,65%	91,54%	95,68%	-	-	93,95%
	Side	83,12%	86,28%	93,52%	-	-	87,64%
	Other	88,65%	93,99%	95,58%	-	-	92,74%
	Total	90,71%	92,35%	94,29%	-	-	92,45%
Sedan	Front	98,10%	97,99%	97,98%	98,40%	96,60%	97,81%
	Back	98,00%	95,97%	96,52%	98,45%	91,27%	96,04%
	Side	94,97%	94,35%	96,25%	97,98%	92,43%	95,20%
	Other	95,83%	97,53%	97,67%	97,84%	86,31%	95,04%
	Total	96,73%	96,46%	97,10%	98,17%	91,65%	96,02%
SUV	Front	95,16%	95,57%	93,43%	92,37%	-	94,13%
	Back	97,67%	95,86%	92,67%	98,88%	-	96,27%
	Side	95,35%	91,30%	90,75%	95,52%	-	93,23%
	Other	92,56%	94,13%	88,95%	96,36%	-	93,00%
	Total	95,18%	94,21%	91,45%	95,78%	-	94,16%
Total	Front	96,11%					94,55%
	Back	95,24%					
	Side	92,91%					
	Other	93,93%					

Tabela 1 – Eficiência da Segmentação

Compact	Back	Front	Other	Side
max_height	659	683	670	624
med_height	580	632	593	554
min_height	487	577	508	491
max_width	644	707	1083	1281
med_width	612	670	1025	1218
min_width	593	641	968	1174
perimetro_max	2606	2780	3506	3810
perimetro_med	2383	2603	3234	3544
perimetro_min	2160	2436	2952	3330
area_max	424396	482881	725610	799344
area_med	354509	423105	607016	674468
area_min	288791	369857	491744	576434

Tabela 2 - Primeiros dados dos Compactos

Pick Up	Back	Front	Other	Side
max_height	838	949	868	770
med_height	773	889	815	729
min_height	669	779	722	662
max_width	1089	1033	1540	1607
med_width	943	917	1395	962
min_width	780	741	1294	1279
perimetro_max	3854	3964	4816	4754
perimetro_med	3432	3611	4420	3381
perimetro_min	2898	3040	4032	3882
area_max	912582	980317	1336720	1237390
area_med	728996	814917	1136704	700735
area_min	521820	577239	934268	846698

Tabela 3 - Primeiros dados das Pick-Up

Sedan	Back	Front	Other	Side
max_height	590	630	606	563
med_height	565	601	581	503
min_height	542	565	563	471
max_width	758	787	1402	1624
med_width	721	728	1204	1497
min_width	672	660	1031	1338
perimetro_max	2696	2834	4016	4374
perimetro_med	2570	2657	3570	4000
perimetro_min	2428	2450	3188	3618
area_max	447220	495810	849612	914312
area_med	406851	437262	699399	753290
area_min	364224	372900	580453	630198

Tabela 4 - Primeiros dados dos Sedan

SUV	Back	Front	Other	Side
max_height	755	731	775	658
med_height	679	684	699	625
min_height	613	651	636	592
max_width	872	852	1417	1706
med_width	775	783	1320	1587
min_width	694	732	1284	1492
perimetro_max	3254	3166	4384	4728
perimetro_med	2908	2932	4037	4425
perimetro_min	2614	2766	3840	4168
area_max	658360	622812	1098175	1122548
area_med	526079	534839	922001	992459
area_min	425422	476532	816624	883264

Tabela 5 - Primeiros dados dos SUVs

	Compact	Sedan	Pick-up	SUV
max_suspensão	75	75	187	124
min_suspensão	59	51	136	95

Tabela 6 - Suspensão dos Veículos

	Compact	Pick-up	Sedan	SUV
max_area	458116	885202	584497	815131
min_area	420298	548875	450796	614303
max_axis	0,5057	0,4119	0,3621	0,4155
min_axis	0,3833	0,3931	0,3116	0,3794
max_formula	25,9889	39,4272	28,0	24,2938
min_formula	23,2623	37,7355	24,7956	22,6734
max_box	770016	1476860	875465	1122548
min_box	624624	1063834	630198	883264

Tabela 7 - Referências/Parâmetros dos Veículos

Relação das alturas	
Sedan & SUV	4,90%
Sedan & Pick-Up	14,95%
SUV & Pick-Up	0,60%
compacto com grande amplitude	

Relação das boxs	
Compacto & SUV	12,82%
Sedan & SUV	0,88%
Sedan & Pick-Up	17,71%

Relação das larguras	
Compacto & SUV	14,14%
Compacto & Sedan	4,26%
Não separa Sedan de SUV	

Relação das áreas	
Compacto & SUV	25,43%
Sedan & SUV	4,85%
Meio termo - Sedan & SUV	15,73%
Pick-Ups com grande amplitude	
Não separa Compacto de Sedan	

Tabela 8 - Margens de Decisão 1

Relação das fórmulas	
Sedan & Pick-Up	25,80%
SUV & Sedan	2,02%
SUV & Pick-Up	35,62%

Relação dos axis	
Sedan & Compacto	5,53%
Sedan & SUV	4,56%
Sedan & Pick-Up	7,89%
Não separa SUV de Compacto	

Tabela 9 - Margens de Decisão 2

	1	2	3	4
Algoritmo 1	0,88%	35,62%	5,53%	
Algoritmo 2	25,80%	4,85%	5,53%	
Algoritmo 3	25,80%	15,73%	4,90%	5,53%
Algoritmo 4	25,80%	4,56%	25,43%	

Tabela 10 - Margens de Decisão associadas aos algoritmos consoante a sua fase na árvore de decisão

Com a Vista Lateral da Segmentação da task 1						Total
Eficiência		Número de Classificações Corretas				
		Pick-Up	Compact	Sedan	SUV	
Algoritmo 1	88%	1	4	5	4	16
Algoritmo 2	81%	0	4	5	4	
Algoritmo 3	81%	0	4	5	4	
Algoritmo 4	81%	0	4	5	4	

Com a Vista Lateral do Ground-Truth						Total
Eficiência		Número de Classificações Corretas				
		Pick-Up	Compact	Sedan	SUV	
Algoritmo 1	100%	3	4	5	4	16
Algoritmo 2	100%	3	4	5	4	
Algoritmo 3	100%	3	4	5	4	
Algoritmo 4	100%	3	4	5	4	

Tabela 11 - Eficiência da Classificação

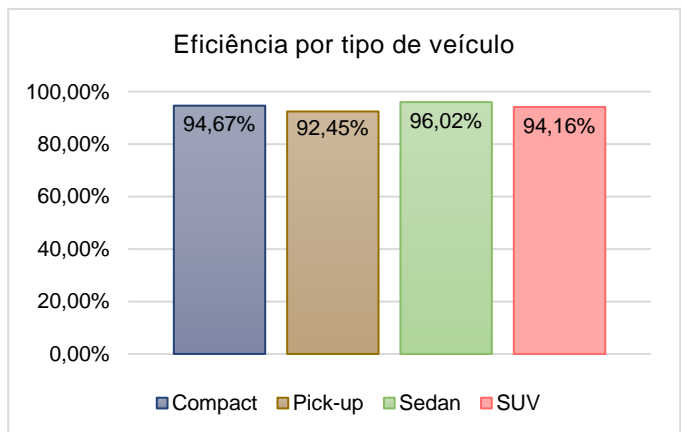


Gráfico 1 - Eficiência por tipo de Veículo

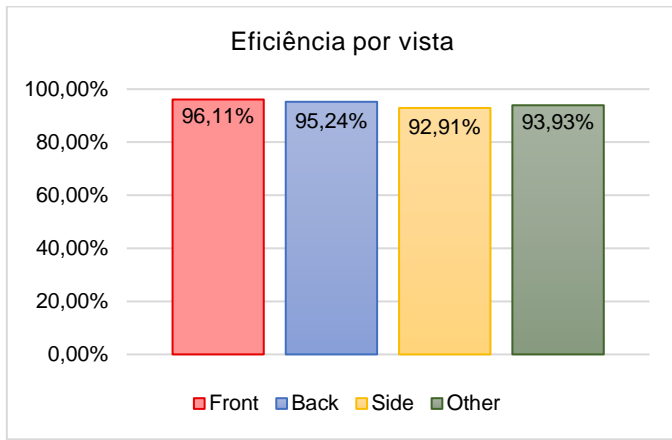


Gráfico 2 - Eficiência por Vista

REFERÊNCIAS

Rafael C.Gonzalez, Richard E.Woods, Steven L. Eddins, *Digital Image Processing Using Matlab*. New Jersey, United States Of America, 2004