

6 Avaliação Experimental

Neste capítulo se discutem os resultados coletados na pesquisa, obtidos com as considerações descritas no capítulo anterior, para o qual foram avaliados os desempenhos dos métodos de otimização de forma quantitativa e visual.

6.1. Avaliação Quantitativa

A avaliação quantitativa baseia-se nos valores do P_{value} e CI calculados pelo teste de hipótese sobre as diferenças de valores de vascularidade produzidos por cada método, conforme descrito no capítulo anterior, assim, o teste de hipótese é aplicado à diferença nos valores de vascularidade de dois métodos, m_o e m_i .

Na Tabela 6 se apresentam os valores do P_{value} obtidos no teste, onde em uma cor verde é representado o método que teve a melhor média da vascularidade avaliada para cada dada imagem de teste é caracterizada também pelo símbolo m_o . As células sombreadas numa cor azul correspondem à situação em que a hipótese nula não foi rejeitada, isto é, no caso em que o teste não encontrou evidência para se supuser que os métodos m_o e m_i apresentem desempenhos distintos, com o que estes métodos teriam um desempenho similar do que aquele com a melhor média da vascularidade em cada imagem.

As demais células da Tabela 6, aquelas que não estão sombreadas por cor alguma, têm a característica do que seu P_{value} é menos do que o nível de significância estabelecido, com isto a hipótese nula foi rejeitada, assim, estes resultados indicam que estes métodos terão um desempenho diferente do que aquele método representado pela média de vascularidade m_o .

Entenda-se por desempenho diferente, à característica do método avaliado de ter um melhor o pior desempenho comparado com o método com a menor média da vascularidade, mas esta definição só poderá ser realizada levando em consideração os valores dos intervalos de confiança para cada método, resultados que são apresentados na Tabela 7.

Exame de CT	Métodos Analisados com o Teste de Hipótese						
	DE	GPS	GPS+ UD	GPS+ DE	MADS	MADS +UD	NM
Synthetic1	0.2 %	0.0 %	49.2 %	20.8 %	0.0 %	4.2 %	m_o
Synthetic2	2.5 %	5.6 %	28.6 %	52.5 %	0.0 %	3.9 %	m_o
Synthetic3	2.5 %	0.0 %	65.9 %	38.0 %	0.0 %	0.0 %	m_o
OPF1	0.8 %	1.2 %	7.6 %	6.3 %	0.0 %	0.0 %	m_o
OPF3	m_o	77.2 %	26.3 %	9.9 %	0.0 %	16.3 %	11.7 %
OPF4	0.0 %	0.1 %	13.4 %	m_o	0.0 %	0.0 %	0.0 %
OPF5	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	m_o
OPF6	0.8 %	1.3 %	12.6 %	10.4 %	0.0 %	10.4 %	m_o
OPF7	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	m_o
OPF8	0.0 %	0.7 %	m_o	4.3 %	0.0 %	0.0 %	6.1 %
OPF9	0.0 %	0.9 %	34.6 %	m_o	0.0 %	0.0 %	87.2 %
Coronary1	1.1 %	0.0 %	72.1 %	m_o	0.0 %	0.3 %	66.0 %
Coronary2	0.0 %	30.4 %	13.0 %	2.2 %	0.0 %	0.4 %	m_o
Liver1	m_o	0.3 %	18.4 %	39.9 %	0.1 %	2.8 %	2.0 %
Liver2	m_o	0.5 %	49.2 %	76.7 %	0.1 %	1.7 %	83.9 %

Tabela 6. Valores do P_{value} para o teste de hipótese¹⁰.

Cabe lembrar que por definição no capítulo anterior, o desempenho de um método cujo intervalo de confiança inclui o valor zero, não pode ser considerado diferente de m_o . De maneira complementar, se todo o intervalo de confiança estiver abaixo de zero, o método em questão deverá ser considerado pior do que o método com a menor média da vascularidade (melhor método), isto é, aquele que apresentou média igual a m_o . De modo análogo, se o intervalo de confiança estiver totalmente acima de zero, o método em questão deverá ser considerado como melhor do que aquele cuja média foi m_o .

Na Tabela 7 são apresentados os valores dos intervalos de confiança para cada um dos métodos avaliados, onde as células sombreadas numa cor verde não possuem valor do CI porque elas representam ao método com o menor valor da média da vascularidade; as células sombreadas numa cor azul representam aos métodos com o desempenho similar ao melhor método encontrado. Por último, as células sem cor são aquelas onde os métodos de

¹⁰ Nesta tabela a referência m_o indica que o método correspondente forneceu o menor valor da vascularidade na média dos 50 pontos avaliados em cada imagem.

otimização apresentaram desempenhos diferentes ao método com o valor da vascularidade m_o .

Imagem de CT	Intervalo de Confiança - CI						
	DE	GPS	GPS+ UD	GPS+ DE	MADS	MADS +UD	NM
Synthetic1	-0,072 -0,017	-0,171 -0,075	-0,037 0,018	-0,054 0,012	-0,209 -0,115	-0,065 -0,001	m_o
Synthetic2	-0,057 -0,004	-0,073 0,001	-0,039 0,012	-0,044 0,023	-0,201 -0,108	-0,069 -0,002	m_o
Synthetic3	-0,036 -0,003	-0,106 -0,038	-0,016 0,010	-0,023 0,009	-0,126 -0,060	-0,106 -0,049	m_o
OPF1	-0,078 -0,013	-0,078 -0,010	-0,066 0,003	-0,063 0,002	-0,103 -0,032	-0,101 -0,036	m_o
OPF3	m_o	-0,016 0,012	-0,005 0,019	-0,002 0,023	-0,044 -0,016	-0,023 0,004	-0,006 0,052
OPF4	-0,022 -0,008	-0,049 -0,013	-0,012 0,002	m_o	-0,052 -0,021	-0,043 -0,015	0,032 0,098
OPF5	-0,262 -0,177	-0,273 -0,184	-0,254 -0,170	-0,245 -0,160	-0,310 -0,229	-0,273 -0,186	m_o
OPF6	-0,110 -0,018	-0,105 -0,013	-0,089 0,011	-0,099 0,010	-0,202 -0,102	-0,094 0,009	m_o
OPF7	-0,156 -0,074	-0,142 -0,049	-0,131 -0,046	-0,165 -0,067	-0,187 -0,100	-0,153 -0,063	m_o
OPF8	-0,069 -0,039	-0,076 -0,013	m_o	-0,044 -0,001	-0,107 -0,050	-0,068 -0,027	-0,002 0,076
OPF9	-0,046 -0,016	-0,050 -0,007	-0,024 0,009	m_o	-0,081 -0,029	-0,065 -0,029	-0,024 0,029
Coronary1	-0,038 -0,005	-0,078 -0,023	-0,023 0,016	m_o	-0,122 -0,074	-0,057 -0,012	-0,044 0,028
Coronary2	-0,071 -0,026	-0,028 0,009	-0,037 0,005	-0,046 -0,004	-0,097 -0,039	-0,050 -0,010	m_o
Liver1	m_o	-0,076 -0,017	-0,029 0,006	-0,024 0,010	-0,083 -0,025	-0,036 -0,002	-0,100 -0,009
Liver2	m_o	-0,065 -0,012	-0,007 0,015	-0,012 0,009	-0,064 -0,019	-0,022 -0,002	-0,041 0,034

Tabela 7. Valor dos CI nos testes de hipótese realizados.

Dos resultados apresentados na Tabela 6 e na Tabela 7, podemos extrair as seguintes conclusões:

- O P_{value} do método *DE* foi sempre menor do que γ , o que indica que seu desempenho foi em todos os exames diferente a do obtido pelo método m_o . Por outro lado, seus intervalos de confiança *CI* estiveram sempre abaixo de zero, indicando, assim, que seu desempenho não será tão bom quanto ao do método m_o .
- Os métodos *GPS*, *MADS* e *MADS+UD*, não produziram a melhor média em nenhum dos casos, isto é, não foram considerados como m_o . Analisando o P_{value} e *CI* verifica-se que o desempenho destes métodos foi quase sempre inferior ao de m_o .
- Na maioria dos testes, os métodos *GPS+UD* e *GPS+DE* tiveram um desempenho igual a m_o . Nos casos em que isso não ocorreu os intervalos de confiança estão muito próximo de zero, indicando que seu desempenho será um pouco pior só. A única exceção foi o método *GPS+UD* aplicado à imagem OPF5 para a qual o seu desempenho foi realmente pior. Observa-se também que cada um dos métodos *DE* e *GPS+DE* foram considerados os melhores em 20% dos casos.
- Por último, o método *NM* apresentou a melhor média em 53.7% dos casos (8 exames). Analisando os intervalos de confiança deste método, nota-se que em 40% dos casos (6 exames) o desempenho de *NM* foi igual ou melhor do que m_o ; num único caso, para o exame Liver1, o intervalo de confiança ficou todo abaixo de zero, indicando que seu desempenho é efetivamente pior.

Adicionalmente aos resultados do teste de hipótese, os dados da Tabela 8 representam a média do número de vezes ¹¹ que a função objetivo foi executada em cada método para minimizar o valor da vascularidade em cada um dos 50 pontos escolhidos de cada exame de CT. Desta tabela podemos extrair as seguintes conclusões:

- O método *DE* executou o número máximo de avaliações da função objetivo permitidas em todos os exames avaliados; este número foi um parâmetro de entrada em nossos experimentos, estabelecido em 400 iterações (vide Tabela 3).

¹¹ Denominada também como iterações, o término faz referência ao número de vezes que a função objetivo é avaliada por um método para minimizar a função objetivo.

- Os métodos *GPS*, *GPS+UD* e *GPS+DE* tiveram um desempenho similar quanto ao número de avaliações; na média estes três métodos envolveram 207 iterações.
- Nos métodos *MADS* e *MADS+UD*, o número de iterações foi praticamente igual à metade do que se observou nos métodos *GPS* e suas variantes, e quase a quarta parte do número de iterações no método *DE*; o número médio de iterações destes métodos foi de 108 iterações.
- O método *NM* foi similar ao método *MADS*, com uma média 112 iterações.

Imagem de CT	Número de Execuções da Função Objetivo nos Métodos de Otimização Implementados						
	DE	GPS	GPS+UD	GPS+DE	MADS	MADS+UD	NM
Synthetic1	400	204	228	204	108	120	124
Synthetic2	400	216	204	210	108	132	123
Synthetic3	400	192	216	192	102	99	127
OPF1	400	228	228	210	123	102	121
OPF3	400	186	204	192	108	105	113
OPF4	400	204	210	198	129	111	124
OPF5	400	180	180	192	84	90	64
OPF6	400	246	216	216	96	123	112
OPF7	400	222	216	192	105	114	97
OPF8	400	228	240	222	114	108	91
OPF9	400	204	204	216	105	84	108
Coronary1	400	222	216	222	114	126	133
Coronary2	400	252	228	228	138	114	127
Liver1	400	168	168	150	84	84	112
Liver2	400	216	180	180	96	102	108

Tabela 8. Número de execuções dos métodos estudados.

Associando as conclusões derivadas da análise quanto à medida de vascularidade e quanto à eficiência computacional, conclui-se que o método *NM* quase sempre obteve o melhor compromisso entre acurácia e desempenho computacional.

Nota-se ainda, que o método *DE* obteve a pior relação entre acurácia e carga computacional. Os métodos *MADS* e *MADS+UD*, embora tenham sido os

mais rápidos na maioria dos casos, não foram estáveis quanto ao critério do cálculo da vascularidade.

Por fim, os resultados permitem recomendar os métodos *GPS+UD* e *GPS+DE* como alternativas viáveis para a otimização da vascularidade, já que apresentaram resultados comparáveis tanto em desempenho como em custo computacional com o método *NM*.

6.2. Avaliação Visual

De maneira a complementar a análise quantitativa, nesta seção se apresentam alguns resultados visuais (Figura 34 a Figura 39) obtidos com a aplicação dos métodos de otimização com o melhor e pior desempenho, *NM* e *DE* respectivamente, sobre os exames de CT completos (apresentados nas Figura 27 à Figura 32) e não só sobre os 50 pontos utilizados para a avaliação quantitativa, isto a fim de observar a importância da escolha de um método de otimização adequado para a determinação dos parâmetros que modelam a estrutura da rede vascular.

No lado esquerdo das imagens é apresentado o modelo da estrutura dos vasos sanguíneos obtido com o método de otimização *DE* e no lado direito apresenta-se o modelo obtido com a aplicação do método de otimização *NM*. Para o caso das Figuras 34, 35 e 36, correspondentes aos resultados das estruturas vasculares nas imagens sintéticas, em uma cor verde é representada a imagem de referência e em uma cor roxa é representado o resultado obtido pelos métodos de otimização.

Estes resultados visuais são coerentes com o modelo 3D de referência, e demonstram que o método de otimização *NM* conseguiu caracterizar adequadamente a estrutura dos vasos.

Assim, na maioria dos casos apresentados a rede vascular principal consegue ser adequadamente segmentada, a partir da definição de uma só semente.

O principal problema que a maioria dos métodos de segmentação da rede vascular apresenta reside na identificação das bifurcações presentes na rede. Embora o método de otimização ajude na segmentação da rede completa através da melhor definição da direção dos cilindros que modelam o vaso, nota-se que o desempenho do algoritmo se degrada próximo a bifurcações.

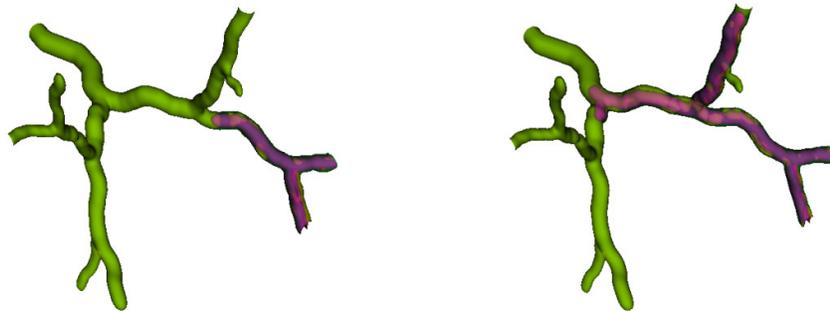


Figura 34. Aplicação do método *DE* (esquerda) e *NM* (direita) no exame Synthetic1.

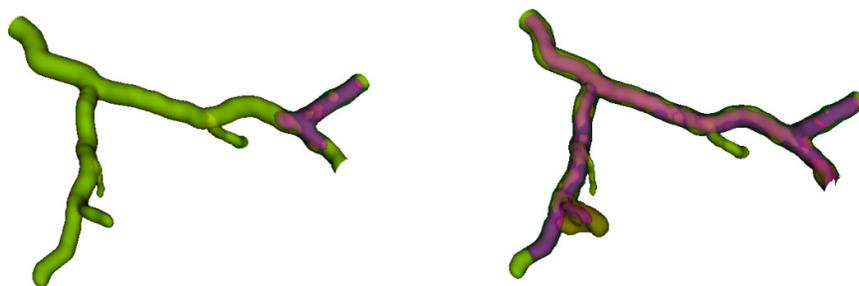


Figura 35. Aplicação do método *DE* (esquerda) e *NM* (direita) no exame Synthetic2.

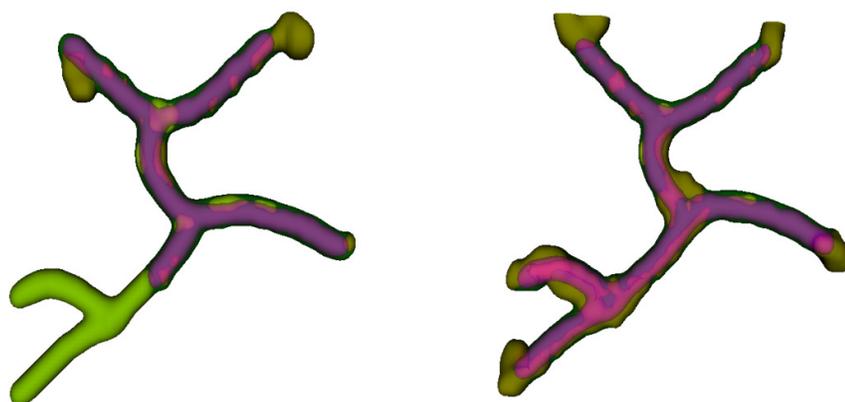


Figura 36. Aplicação do método *DE* (esquerda) e *NM* (direita) no exame Synthetic3.



Figura 37. Aplicação do método *DE* (esquerda) e *NM* (direita) no exame OPF1.



Figura 38. Aplicação do método *DE* (esquerda) e *NM* (direita) no exame OPF3.



Figura 39. Aplicação do método *DE* (esquerda) e *NM* (direita) no exame OPF7.