



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA
MESTRADO ACADÊMICO EM ENGENHARIA DE TELEINFORMÁTICA

PAULO CIRILLO SOUZA BARBOSA

PARAMETRIZAÇÃO DE SINAIS DE ELETROENCEFALOGRAMA PARA
CLASSIFICAÇÃO DE PADRÕES VIA MATRIZES DE KERNEL

FORTALEZA

2021

PAULO CIRILLO SOUZA BARBOSA

PARAMETRIZAÇÃO DE SINAIS DE ELETROENCEFALOGRAMA PARA
CLASSIFICAÇÃO DE PADRÕES VIA MATRIZES DE KERNEL

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Orientador: Prof. Dr. Guilherme de Alencar Barreto

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- B211p Barbosa, Paulo.
Parametrização de sinais de eletroencefalograma para classificação de padrões via matrizes de kernel / Paulo Barbosa. – 2021.
155 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2021.
Orientação: Prof. Dr. Guilherme de Alencar Barreto.
1. Eletroencefalograma. 2. Parametrização de sinais. 3. Matrizes de kernel. 4. Esparsificação de modelos.
I. Título.

CDD 621.38

PAULO CIRILLO SOUZA BARBOSA

PARAMETRIZAÇÃO DE SINAIS DE ELETROENCEFALOGRAMA PARA
CLASSIFICAÇÃO DE PADRÕES VIA MATRIZES DE KERNEL

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Aprovada em:

BANCA EXAMINADORA

Prof. Dr. Guilherme de Alencar Barreto (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Frederico Gadelha Guimarães (PPGEE)
Universidade Federal de Minas Gerais (UFMG)

Prof. Dr. José Maria Pires de Menezes Júnior
(PPGEE)
Universidade Federal do Piauí (UFPI)

Prof. Dr. André Lima Férrer de Almeida (PPGETI)
Universidade Federal do Ceará (UFC)

Dedico este trabalho aos queridos familiares que
confiam sempre nos meus esforços.

AGRADECIMENTOS

Agradeço ao meu orientador, Prof. Dr. Guilherme de Alencar Barreto, pela confiança, paciência e ensinamentos, durante o desenvolvimento da presente dissertação e outros trabalhos ao longo do curso de mestrado.

À minha mãe, Lise Mary Soares Souza, que me criou sempre com amor e carinho e nunca descreditou no meu potencial.

Agradeço ao meu pai, Francisco de Assis Barbosa Filho que proveu sempre, os meios necessários para meus estudos.

Aos meus avós, José Maria Rodrigues Souza e Lúcia Soares Souza, que me educaram, cuidaram, e sempre foram grandes apoiadores de todas as minhas batalhas.

À minha namorada, Sharon Karolinne Melo que está sempre ao meu lado, me apoiando e me motivando.

Agradeço aos colegas de pós-graduação do Departamento de Engenharia de Teleinformática (DETI), Raul Fontenele, Renan Fonteles, Renan Bessa, David Nascimento e Polycarpo Souza, por todas as trocas de experiências e auxílios durante o curso.

À Fundação Cearense de Apoio ao Desenvolvimento (Funcap), pela concessão da bolsa de estudos referente a pesquisa do presente trabalho.

“The art of doing mathematics is finding that special case that contains all the germs of generality.”

(David Hilbert)

RESUMO

Eletroencefalogramas (EEGs) fazem parte de uma categoria de sinais biopotenciais amplamente utilizados para auxílio em diagnósticos relacionados a distúrbios neurológicos. Entretanto, estes sinais possuem características complexas, que comumente acarretam diagnósticos errôneos. Por isso, sinais de EEG são objeto de estudo em pesquisas que buscam um melhor entendimento de tais características e suas correlações com determinadas patologias e estados mentais. Isto posto, esta dissertação trata do processo de extração de atributos nos sinais de EEG para fins de classificação. Esta tarefa, também conhecida como parametrização do sinal, é amplamente utilizada na área de aprendizado de máquinas e possui diversos métodos clássicos para extrair informações úteis do sinal. Neste contexto, o presente trabalho introduz novas estratégias de parametrização dos sinais EEG, a partir do uso de matrizes de kernel para construção de vetores de atributos rotulados. Para tal, utilizam-se métodos de esparsificação de modo a reduzir a dimensão das matrizes de kernel. A construção do vetor de atributos segue duas diferentes estratégias: (i) modo intercanais, em que os sinais de EEG dos vários canais são processados conjuntamente; e (ii) modo intracanal, em que os canais são processados individualmente. Os métodos propostos na presente dissertação são comparados com dois métodos lineares clássicos, a saber: vetor de covariâncias e vetor de coeficientes de codificação linear preditiva (*linear predictive coding*, LPC). Utilizam-se apenas discriminantes lineares para a tarefa de classificação, uma vez que o objetivo é avaliar a qualidade dos métodos de extração e não realizar um amplo estudo comparativo entre classificadores. Uma extensa avaliação de desempenho dos métodos propostos a partir da variação dos seus hiperparâmetros é executada. Para isso, dois conjuntos de dados são utilizados para gerar os resultados do presente trabalho. A partir da análise e discussão dos resultados obtidos, foi possível verificar que os métodos propostos são promissores, apresentando desempenho equivalente ou superior aos gerados pelos métodos clássicos de extração de atributos de sinais de EEG.

Palavras-chave: Eletroencefalograma. Parametrização de sinais. Matrizes de kernel. Esparsificação de modelos.

ABSTRACT

Electroencephalograms (EEGs) belong to a category of biopotential signals widely used in healthcare for the diagnosis of neurological disorders. However, these signals have complex characteristics that commonly lead to misdiagnosis. Hence, EEG signals are the object of study in researches that seek a better understanding of those characteristics and their correlations with certain pathologies and mental states. That said, the present master thesis addresses the process of extracting attributes from EEG signals for classification purposes. This task, also known as signal parameterization, is widely used in machine learning and has several classical methods to extract useful information from the signal. In this context, the present work introduces new EEG signal parameterization strategies, from the use of kernel matrices to build labeled attribute vectors. For that, sparsification methods are used in order to reduce the dimension of the kernel matrices. The construction of the attribute vector follows two different strategies: (i) interchannel mode, in which the EEG signals of the various channels are processed together; and (ii) intrachannel mode, in which channels are processed individually. The methods proposed in this master thesis are compared with two classical linear methods, namely: vector of covariances and vector of linear predictive coding coefficients (LPCC). Only linear discriminants are used for the classification task, since the objective is to assess the quality of the extraction methods and not to evaluate classifiers per se. A comprehensive comparative study evaluating the performance of the proposed methods from the variation of their hyperparameters is carried out. For this purpose, two benchmarking data sets are used to generate the results of the present work. From the analysis and discussion of the results obtained, it was possible to infer that the proposed methods are promising, with performance equivalent or superior to those generated by the classical methods of attribute extraction of EEG signals.

Keywords: Electroencephalogram. Signal parametrization. Kernel matrices. Model sparsification.

LISTA DE FIGURAS

Figura 1 – Comportamento do potencial de ação devido ao estímulo.	26
Figura 2 – Representação neuronal e seus principais grupos.	27
Figura 3 – Esquemático do procedimento de aquisição de EEG no couro cabeludo. . .	28
Figura 4 – Sinal Eletroencefalograma (EEG) registrado por Hans Berger.	29
Figura 5 – Eletrodos comumente utilizados na Aquisição EEG: a) Eletrodos de ponte; b) Eletrodo de copo; c) Touca com eletrodos já fixados.	30
Figura 6 – Exemplo do sistema internacional 10 – 20.	32
Figura 7 – Sistema completo para eletroencefalografia.	33
Figura 8 – a) Sinal lento monomórfico; b) Sinal lento polimórfico; c) Sinal agudo. . . .	34
Figura 9 – Sistema completo para eletroencefalografia.	35
Figura 10 – Exemplos de sinais por faixa de frequência.	37
Figura 11 – Artefatos por movimento ocular.	38
Figura 12 – Artefato causado pela interferência da rede elétrica.	39
Figura 13 – Abordagem intracanal.	41
Figura 14 – Abordagem intercanais.	41
Figura 15 – Fluxograma da parametrização via LPC de uma época.	43
Figura 16 – Fluxograma da parametrização via covariância pelo modo intercanais. . . .	44
Figura 17 – Fluxograma da parametrização via covariância pelo modo intracanal. . . .	46
Figura 18 – Exemplo de um mapeamento para espaço de características.	48
Figura 19 – Fluxograma de parametrização pela abordagem intercanais em uma época. .	56
Figura 20 – Fluxograma de parametrização pela abordagem intracanal em uma época. .	57
Figura 21 – Padrão 10-20 utilizado no conjunto de dados.	63
Figura 22 – Boxplot com todas as quantidades de períodos convulsivos.	63
Figura 23 – Segmentações utilizadas para a etapa de treinamento.	64
Figura 24 – Sistema 10-20 de posicionamento e escolha de canais.	65
Figura 25 – Exemplo de sinal para predisposição ao alcoolismo e grupo de controle. . .	66
Figura 26 – Fluxograma geral para armazenamento de matrizes parametrizadas no pri- meiro conjunto de dados.	67
Figura 27 – Fluxograma geral para armazenamento de matrizes parametrizadas no se- gundo conjunto de dados.	68
Figura 28 – Matrizes geradas no processo de parametrização a serem armazenadas. . .	68

Figura 29 – Processo de validação cruzada em j dobras.	74
Figura 30 – Agrupamento dos grupos de matrizes parametrizadas do segundo conjunto de dados.	75
Figura 31 – Processo de treino e teste para segundo conjunto de dados.	75
Figura 32 – Exemplo de Matriz de Confusão.	76
Figura 33 – Comparação entre <i>Boxplot</i> e <i>Violin plot</i>	78
Figura 34 – Exploração dos hiperparâmetros da função kernel polinomial.	80
Figura 35 – Variações de p com aplicação no kernel polinomial.	81
Figura 36 – Exploração kernel RBF pela abordagem intercanais.	81
Figura 37 – Exploração kernel RBF e abordagem intracanal.	82
Figura 38 – Exploração de p em modo intracanal para kernel RBF com esparsificação pela entropia de Renyi.	83
Figura 39 – CHB01 : Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.	84
Figura 40 – CHB01 : Comparativo das sensibilidades entre OLS e CLG para segmentação 01.	84
Figura 41 – CHB01 : Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.	85
Figura 42 – CHB01 : Comparativo das sensibilidades entre OLS e CLG para segmentação 02.	86
Figura 43 – CHB01 Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.	87
Figura 44 – CHB01 Comparativo das sensibilidades entre OLS e CLG para segmentação 03.	87
Figura 45 – CHB01 : Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.	88
Figura 46 – CHB01 : Comparativo das sensibilidades entre OLS e CLG para segmentação 04.	89
Figura 47 – CHB03 Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 01.	90
Figura 48 – CHB03 : Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.	91

Figura 49 – CHB03 Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 02.	92
Figura 50 – CHB03 Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.	93
Figura 51 – CHB03: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 03.	94
Figura 52 – CHB03 Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.	95
Figura 53 – CHB03: Comparativo das sensibilidades entre OLS e CLG para segmentação 04.	95
Figura 54 – CHB03: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.	96
Figura 55 – CHB14: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 01.	97
Figura 56 – CHB14 Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.	98
Figura 57 – CHB14: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 02.	98
Figura 58 – CHB14: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.	99
Figura 59 – CHB14: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 03.	100
Figura 60 – CHB14: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.	101
Figura 61 – CHB14: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 04.	102
Figura 62 – CHB14: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.	102
Figura 63 – CHB18 Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 01.	103
Figura 64 – CHB18 Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.	104

Figura 65 – CHB18: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 02.	105
Figura 66 – CHB18: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.	106
Figura 67 – CHB18: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 03.	106
Figura 68 – CHB18: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.	107
Figura 69 – CHB18: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 04.	108
Figura 70 – CHB18: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.	108
Figura 71 – CHB24: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 01.	109
Figura 72 – CHB24: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.	110
Figura 73 – CHB24: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 02.	111
Figura 74 – CHB24 Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.	111
Figura 75 – Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 03.	112
Figura 76 – CHB24: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.	113
Figura 77 – CHB24: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 04.	114
Figura 78 – CHB24: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.	114
Figura 79 – Exploração k Kernel RBF intracanal.	115
Figura 80 – Exploração m Kernel RBF intracanal.	116
Figura 81 – Exploração $m = k$ Kernel polinomial intracanal.	117
Figura 82 – Métricas obtidas pelo classificador OLS.	118

Figura 83 – Taxas de sensibilidade, especificidade e acurácia obtidas pelo CLG.	120
Figura 84 – Custo computacional métodos no primeiro conjunto de dados.	122
Figura 85 – Custo computacional métodos no segundo conjunto de dados.	122
Figura 86 – Resultados do caso CHB02	138
Figura 87 – Resultados do caso CHB04	139
Figura 88 – Resultados do caso CHB05	140
Figura 89 – Resultados do caso CHB06	141
Figura 90 – Resultados do caso CHB07	142
Figura 91 – Resultados do caso CHB08	143
Figura 92 – Resultados do caso CHB09	144
Figura 93 – Resultados do caso CHB10	145
Figura 94 – Resultados do caso CHB11	146
Figura 95 – Resultados do caso CHB12	147
Figura 96 – Resultados do caso CHB13	148
Figura 97 – Resultados do caso CHB16	149
Figura 98 – Resultados do caso CHB17	150
Figura 99 – Resultados do caso CHB19	151
Figura 100–Resultados do caso CHB20	152
Figura 101–Resultados do caso CHB21	153
Figura 102–Resultados do caso CHB22	154
Figura 103–Resultados do caso CHB23	155

LISTA DE TABELAS

Tabela 1 – Faixas de Frequência de sinais EEG.	35
Tabela 2 – Nomenclaturas utilizadas nos gráficos dos resultados.	79
Tabela 3 – Resultados médios ao utilizar o OLS.	119
Tabela 4 – Resultados médios ao utilizar Classificador Linear Gaussiano (CLG).	120
Tabela 5 – Melhores resultados para o caso CHB01	133
Tabela 6 – Melhores resultados para o caso CHB03	134
Tabela 7 – Melhores resultados para o caso CHB14	135
Tabela 8 – Melhores resultados para o caso CHB18	136
Tabela 9 – Melhores resultados para o caso CHB24	137

LISTA DE ALGORITMOS

Algoritmo 1	– k -Médias.	53
Algoritmo 2	– Esparsificação pela entropia.	55
Algoritmo 3	– Metodologia de parametrização via <i>Linear Predictive Coding</i> (LPC). . .	69
Algoritmo 4	– Metodologia de parametrização via matriz de Covariância (intercanais). . .	70
Algoritmo 5	– Metodologia de parametrização via matriz de Covariância (Intracanal). . .	71
Algoritmo 6	– Metodologia de parametrização via matriz de Kernel (Intercanais). . . .	72
Algoritmo 7	– Metodologia de parametrização via matriz de Kernel (Intracanal). . . .	73

LISTA DE ABREVIATURAS E SIGLAS

EEG	Eletroencefalograma
CLG	Classificador Linear Gaussiano
LPC	<i>Linear Predictive Coding</i>
AM	Aprendizado de Máquina
EMG	Eletromiograma
ECG	Eletrocardiograma
EPSP	<i>Excitatory PostSynaptic Potential</i>
IPSP	<i>Inhibitory PostSynaptic Potential</i>
IFSECN	<i>International Federation of Societies for EEG and Clinical Neurophysiology</i>
EOG	Eletrooculograma
AR	Autorregressivo
SVM	<i>support vector machines</i>
RBF	<i>radial basis function</i>
OLS	<i>ordinary least squares</i>
VP	Verdadeiro Positivo
VN	Verdadeiro Negativo
FN	Falso Negativo
FP	Falso Positivo

LISTA DE SÍMBOLOS

<i>Hz</i>	Hertz
<i>K</i>	Quantidade de canais do eletroencefalógrafo
<i>ms</i>	Milissegundos
<i>V</i>	Volts
<i>mV</i>	Milivolts
<i>K⁺</i>	Íons positivos de Potássio
<i>Na⁺</i>	Sódio
<i>mmol/l</i>	Milimol por litro
<i>Ag/AgCl</i>	Prata/Cloreto de prata

SUMÁRIO

1	INTRODUÇÃO	20
1.1	Justificativa	21
1.2	Objetivos	22
1.2.1	<i>Objetivo Geral</i>	22
1.2.2	<i>Objetivos Específicos</i>	22
1.3	Produção Científica	23
1.4	Organização da Dissertação	24
2	FUNDAMENTOS DE ELETROENCEFALOGRAFIA	25
2.1	Sinais Biopotenciais	25
2.2	Método de Eletroencefalografia	28
2.3	Características de sinais EEG	33
2.4	Considerações Finais	39
3	EXTRAÇÃO DE ATRIBUTOS DE SINAIS DE EEG: MÉTODOS CLÁSSICOS	40
3.1	Vetor de Coeficientes LPC	42
3.2	Vetor de Covariâncias	43
3.3	Considerações Finais	45
4	EXTRAÇÃO DE ATRIBUTOS DE SINAIS DE EEG: MÉTODOS PROPOSTOS	47
4.1	Extração através de matrizes de kernel	47
4.2	Esparsificação de Modelos	52
4.2.1	<i>Esparsificação através de quantização vetorial</i>	52
4.2.2	<i>Esparsificação pelo critério de entropia de Renyi</i>	54
4.3	Classificadores	56
4.3.1	<i>Método dos mínimos quadrados</i>	57
4.3.2	<i>Classificador Linear Gaussiano</i>	59
4.4	Considerações Finais	60
5	METODOLOGIA DOS EXPERIMENTOS	62
5.1	Descrição do primeiro conjunto de dados	62
5.2	Descrição do segundo conjunto de dados	64

5.3	Metodologia de Parametrização dos sinais	65
5.3.1	<i>Metodologia de parametrização via LPC</i>	<i>67</i>
5.3.2	<i>Metodologia de parametrização pela matriz de covariância</i>	<i>70</i>
5.3.3	<i>Metodologia de parametrização através dos métodos propostos</i>	<i>70</i>
5.4	Metodologia para treino e teste dos classificadores.	72
5.5	Avaliação de desempenho	75
5.6	Considerações Finais	77
6	RESULTADOS E DISCUSSÕES	79
6.1	Resultados para o primeiro conjunto de dados	80
6.1.1	<i>Resultados caso CHB01</i>	<i>83</i>
6.1.2	<i>Resultados caso CHB03</i>	<i>89</i>
6.1.3	<i>Resultados do caso CHB14</i>	<i>96</i>
6.1.4	<i>Resultados caso CHB18</i>	<i>101</i>
6.1.5	<i>Resultados caso CHB24</i>	<i>109</i>
6.2	Resultados para o segundo conjunto de dados	115
6.3	Avaliação de custo computacional	121
6.4	Considerações Finais	121
7	CONCLUSÕES E TRABALHOS FUTUROS	124
7.1	Trabalhos futuros	125
	REFERÊNCIAS	126
	APÊNDICES	133
	APÊNDICE A–RESULTADOS MÉDIOS CASO CHB01	133
	APÊNDICE B–RESULTADOS MÉDIOS CASO CHB03	134
	APÊNDICE C–RESULTADOS MÉDIOS CASO CHB14	135
	APÊNDICE D–RESULTADOS MÉDIOS CASO CHB18	136
	APÊNDICE E–RESULTADOS MÉDIOS CASO CHB24	137
	APÊNDICE F–DEMAIS RESULTADOS	138

1 INTRODUÇÃO

As últimas décadas do século XXI tem verificado um aumento de pesquisas que utilizam métodos de Aprendizado de Máquina (AM) em diversos setores da sociedade, tais como na economia, publicidade, transporte, indústria manufatureira, saúde, dentre outros (QAYYUM *et al.*, 2021). Tal intensificação é consequência direta dos avanços tecnológicos, que proporcionaram um alto poder computacional, grande capacidade de armazenamento, alta taxa de transmissão e processamento de dados a diversos dispositivos, seja pessoal ou não. (KOCHEUROV *et al.*, 2018; DOUPE *et al.*, 2019).

AM é um ramo da inteligência artificial que utiliza-se de modelos matemáticos e computacionais, com vistas a resolver problemas ou melhorar o desempenho de processos. Tais modelos possuem capacidade de aprendizagem através do reconhecimento de padrões a partir de um conjunto dados (GOLDENBERG *et al.*, 2019; RASHEED *et al.*, 2021). Este é caracterizado como conjunto de treinamento e é usado como um guia para fazer previsões a partir de novas entradas. Comumente, esta etapa de treinamento pode ser dividida em duas categorias: supervisionado; não supervisionado (HOSSEINI *et al.*, 2021). Na primeira, o treinamento é feito a partir de exemplos, ou seja, comparam-se os resultados empíricos com as saídas corretas a fim de identificar erros. Para o não supervisionado, realiza-se uma análise do conjunto de dados sem se conhecer as saídas, a fim de identificar padrões nestes dados (LECUN *et al.*, 2015).

Como dito anteriormente, a utilização de AM está difundida em vários setores, porém pode-se destacar trabalhos em que aplicou-se tais métodos no ramo da medicina, especificamente os que lidaram com sinais biopotenciais, como por exemplo: utilização de AM em sinais de Eletromiograma (EMG) para controle de prótese (KRASOULIS *et al.*, 2020); identificação de arritmia cardíaca em sinais de Eletrocardiograma (ECG) (ALFARAS *et al.*, 2019); diagnóstico de doença de Alzheimer utilizando sinais de EEG e AM (TZIMOURTA *et al.*, 2021) e etc. Esses sinais biopotenciais são objeto de estudo de diversas pesquisas e tratam de medidas elétricas das atividades fisiológicas nas células ou tecidos vivos (CARDOSO, 2010).

O EEG, que é um dos tipos de sinais biopotencial, é adquirido através da eletroencefalografia, sendo um método que grava atividades cerebrais através de eletrodos posicionados no couro cabeludo. Normalmente, esta técnica é utilizada para o monitoramento nos domínios do tempo ou frequência (DELORME; MAKEIG, 2004). O EEG é um sinal elétrico produzido pelo potencial mensurável das células piramidais pós-sinápticas, através da troca de informações de neurônios, a partir de algum tipo de estímulo (SPRUSTON, 2008; SHARMA *et al.*, 2020).

Os sinais de EEG são amplamente utilizados na área clínica, pois auxiliam na identificação e tratamento de doenças como esquizofrenia, Alzheimer, tumores cerebrais, distúrbios de sono, epilepsia, dentre outros (RAMOS-AGUILAR *et al.*, 2020). Além disto, o procedimento de aquisição dos sinais EEG é de baixo custo, e não necessita de grandes restrições de segurança ao aplicar em pacientes (SAVADKOOHI *et al.*, 2020). Apesar de desempenharem um papel vital nas aplicações mencionadas, o processo de análise dos sinais de EEG é complexo e, portanto, demanda uma grande quantidade de tempo e esforços. De acordo com Maszczyk *et al.* (2019), extrair as informações pertinentes de um sinal de EEG através de inspeção visual, é desafiador até mesmo para neurologistas treinados, pelo fato da grande suscetibilidade de ruídos como artefatos ou dos provenientes da rede elétrica.

As principais causas da presença de artefatos no sinal EEG são devido a falhas no dispositivo, eletrodos desacoplados ou não posicionados corretamente, ruídos da rede elétrica e alta impedância nos eletrodos (GE *et al.*, 2017). Além disto, como os sinais de EEG operam em baixas frequências ($< 30Hz$), é comum a presença de outros sinais fisiológicos, como por exemplo: piscadas e movimentos de olhos; atividade cardíaca e atividade muscular. Estes sinais interferem nas informações neurais, influenciando assim a interpretação visual e o diagnóstico em pesquisas clínicas (TAMBURRO *et al.*, 2018; GE *et al.*, 2017).

Assim, em face de tais dificuldades, muitas pesquisas ao longo das últimas décadas foram realizadas de modo a propor novas ferramentas para auxílio ao diagnóstico automatizado no ramo da neurologia (SIJLY; ZHANG, 2016). De acordo com Baig *et al.* (2019), é comum que essas ferramentas sejam compostas das etapas: aquisição do sinal; pré-processamento; extração e seleção de atributos, e por fim classificação. No entanto, há grandes desafios ao trabalhar com os sinais de EEG por possuírem características não lineares, aleatórias e não estacionárias (MO; ZHAO, 2015). Portanto, a etapa de extração de atributos apresenta-se como fundamental para que seja possível extrair informações úteis do sinal de EEG, que então podem ser utilizadas na etapa de classificação (ALÇIN *et al.*, 2016).

1.1 Justificativa

Como dito anteriormente, pelas problemáticas ao se trabalhar com sinais de EEG, há uma necessidade em se desenvolver um sistema que consiga classificar corretamente estes sinais prevenindo que o analista faça interpretações erradas, bem como auxiliando no tratamento de doenças neurológicas (ALÇIN *et al.*, 2016). Destaca-se ainda que a etapa de classificação destes

sinais é realizada adequadamente quando se utilizam métodos de extração de características, já que devem capturar adequada as características dinâmicas dos sinais de EEG. Portanto, esta etapa tem um efeito direto no desempenho da etapa de classificação (MO; ZHAO, 2015).

Há diversos métodos de extração de características utilizados em sinais de EEG. Porém, podem-se destacar dois métodos clássicos que são amplamente utilizados em processamento estatístico de sinais. O primeiro chamado de *linear predictive coding* (LPC), que realiza uma modelagem temporal no sinal e de maneira isolada para cada canal. O segundo, que utiliza as informações da matriz de covariância, faz o uso de abordagem espacial (i.e., não temporal) simultânea para todos os canais. Pode-se destacar que a matriz de covariância reflete similaridades (i.e. correlações lineares) entre variáveis aleatórias.

Com o que foi dito, pode-se ter como base para a presente dissertação, a seguinte hipótese: sendo a matriz de kernel uma matriz gramiana que captura similaridades (i.e., produtos escalares) entre vetores no espaço de reprodução de Hilbert (*reproducing kernel Hilbert space*, RKHS), poderia esta matriz capturar dependências não lineares úteis para o problema de classificação de sinais de EEG?

1.2 Objetivos

Com a descrição do problema e motivação já definidas, nesta seção faz-se a exposição dos objetivos a serem atingidos na presente dissertação.

1.2.1 *Objetivo Geral*

O objetivo geral do presente trabalho é propor uma nova abordagem que faz o uso de matrizes de kernel para extrair informações úteis de segmentos do sinal de EEG (i.e. parametrização de sinal), transformando-as em vetores de atributos rotulados e são utilizados na etapa de treinamento de classificadores. Além disto, propõem-se novas estratégias para construção dessas matrizes pela análise isolada (intra canal) ou conjunta (intercanais) dos canais de EEG.

1.2.2 *Objetivos Específicos*

Os objetivos específicos da presente dissertação estão listados a seguir.

- Conduzir estudo sobre os métodos clássicos de parametrização dos sinais de

EEG.

- Conduzir estudo sobre a composição das matrizes de kernel.
- Realizar a escolha e a implementação de duas funções diferentes para cálculo das matrizes de kernel.
- Realizar a implementação dos algoritmos de esparsificação, a saber, quantização vetorial e critério de maximização de entropia; de modo a diminuir a dimensão das matrizes de kernel.
- Aplicar os métodos propostos através das análises intracanal e intercanais.
- Realizar a implementação dos métodos de classificação. classificador linear de mínimos quadrados (OLS) e classificador linear gaussiano (CLG).
- Escolher dois conjuntos de dados e tentar resolver o problema proposto com todos os métodos implementados.
- Realizar uma exploração dos hiperparâmetros dos métodos de extração de atributos propostos, a fim de identificar os que causam um aumento de desempenho na etapa de classificação.
- Realizar uma análise comparativa entre os métodos propostos e clássicos, através das métricas de acurácia, sensibilidade e especificidade obtidas pela etapa de treinamento.

1.3 Produção Científica

Ao longo do curso de mestrado, foram produzidos os seguintes trabalhos científicos:

1. Barbosa, P., Barreto, G. A. Parametrização de Sinais de EEG para Diagnóstico de Epilepsia Usando Matrizes de Kernel. Trabalho publicado/aceito em anais do 15 Congresso Brasileiro de Inteligência Computacional (CBIC 2021) Joinville, SC. Disponível em <<https://dx.doi.org/10.21528/CBIC2021-118>>.
2. Barbosa, P., A, Santana, R., The, G. Barreto, G. Reconhecimento de expressão facial assistido por EMG e aprendizado de máquina. Trabalho publicado/aceito em anais do 15 Congresso Brasileiro de Inteligência Computacional (CBIC 2021) Joinville, SC. Disponível em: <<https://dx.doi.org/10.21528/CBIC2021-100>>.

1.4 Organização da Dissertação

A organização do restante da presente dissertação é realizada como descrito a seguir.

No Capítulo 2, faz-se a revisão bibliográfica com respeito ao EEG. São descritos os fundamentos sobre os sinais biopotenciais, o método de eletroencefalografia, as características dos sinais EEG.

No Capítulo 3, descreve-se sobre a extração de atributos nos sinais de EEG, realizando uma fundamentação dos métodos clássicos de parametrização via coeficientes LPC e via matriz de covariância.

O Capítulo 4 é destinado aos fundamentos relacionados aos métodos de parametrização propostos pelo presente trabalho. Fundamentam-se os conceitos relacionados as matrizes de kernel e os métodos de esparsificação. É descrito neste capítulo, como os conceitos são utilizados em conjunto, de modo a construir vetores de atributos rotulados. Por fim, na última seção do Capítulo, são expostos os fundamentos com respeito aos classificadores (OLS e CLG) utilizados no trabalho.

No capítulo 5, é descrito sobre a metodologia dos experimentos, utilizada no presente trabalho. Descrevem-se as características dos dois conjuntos de dados usados para aplicar os métodos propostos e clássicos. É explicado através de fluxogramas, como cada um destes métodos são utilizados para parametrizar o sinal de EEG. Explica-se ao final do capítulo, qual é a metodologia para exploração dos hiperparâmetros dos métodos de parametrização.

No capítulo 6, são expostos os resultados obtidos para cada conjunto de dados sendo levantadas discussões com relação aos mesmos.

Por fim, no capítulo 7 são apresentadas as conclusões obtidas com o presente trabalho, bem como um direcionamento para possíveis trabalhos futuros.

2 FUNDAMENTOS DE ELETROENCEFALOGRAFIA

No presente capítulo¹, serão fundamentadas as informações necessárias sobre os sinais de EEG. Inicialmente faz-se uma breve explicação histórica sobre o que são sinais biomédicos, especificamente os de biopotencial e descreve-se sobre as motivações do ser humano ao estudar estes sinais. Em seguida, é explicado sobre as características principais do sinal EEG, bem como dos processos para sua aquisição.

2.1 Sinais Biopotenciais

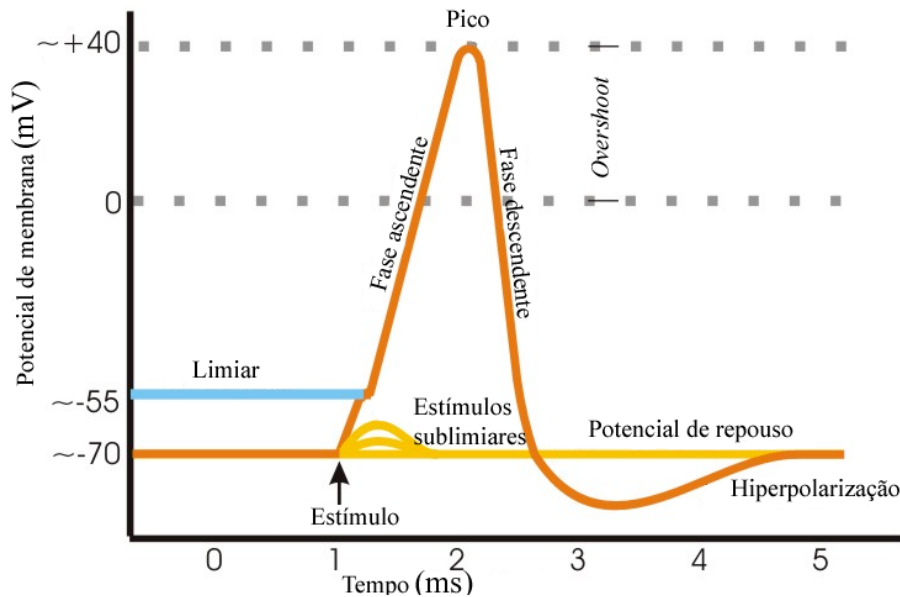
De acordo com Webster (2009) os sinais biopotenciais são provenientes da atividade eletroquímica das células excitáveis presentes nos tecidos neurais, musculares ou glandulares. A atividade dessas células é categorizada como potencial ativo quando está recebendo algum tipo de estímulo, ou potencial em repouso caso contrário. Essa atividade é medida entre a diferença de potencial intra e extracelular (CARDOSO, 2010).

Considerando o caso do potencial de repouso, o comportamento da membrana da célula é caracterizado por ser ligeiramente permeável ao Na^+ e totalmente permeável ao K^+ . Em repouso, a concentração de K^+ é superior no interior da célula do que no exterior. Essa diferença cria um gradiente difusor, fazendo com que os íons de K^+ sejam movimentados para fora da membrana e íons de Na^+ para dentro. Por consequência, a concentração interna se torna cada vez mais negativa com relação ao meio externo e desta maneira, o potencial de repouso é estabelecido quando o equilíbrio é atingido (KANDEL *et al.*, 2000; WEBSTER, 2009).

No entanto, quando se trata dos potenciais ativos em uma célula excitável, há canais chamados de gatilháveis, pois, apenas desempenham sua função de movimentação de íon a partir de determinada situação. Tais canais são muito comuns em células musculares ou neuronais, e realizam seu papel quando há um estímulo físico ou substância química presente (WEBSTER, 2009). O funcionamento destes canais pode ser exemplificado através da Figura 1. Nesta é possível verificar que inicialmente o canal está em repouso e quando um potencial próximo de $-55mV$ é atingido, realiza-se a sua abertura num intervalo pequeno de tempo (aproximadamente $1ms$). Quando isso ocorre, uma grande quantidade de outros canais se abre, modificando novamente a permeabilidade da membrana celular. Sendo assim, com a união dessas atividades

¹ Os fundamentos de eletroencefalografia listados no presente capítulo foram extraídos principalmente de (ME-CARELLI, 2019), (NIEDERMEYER; SILVA, 2020) e (WEBSTER, 2009) e correspondem a um compilado de tópicos lidos ao longo do desenvolvimento da dissertação, e foram organizados para facilitar o entendimento do leitor leigo na área de processamento de bio-sinais.

Figura 1 – Comportamento do potencial de ação devido ao estímulo.



Fonte: Cardoso (2010).

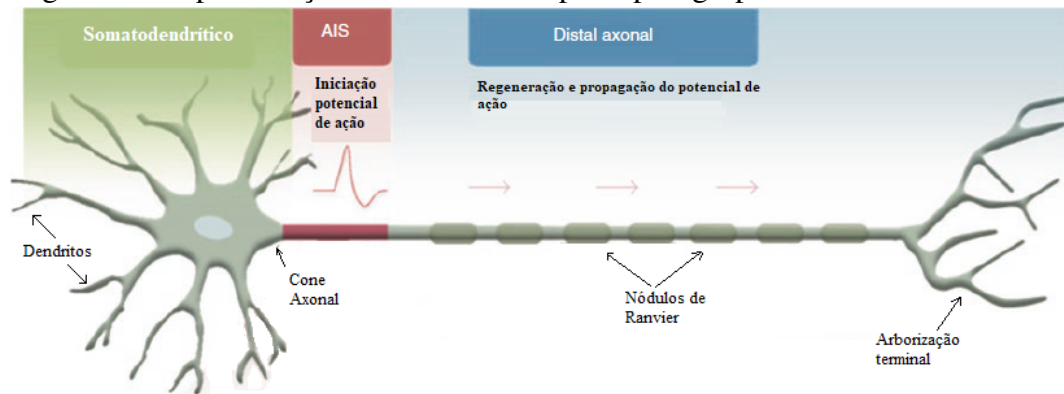
celulares, os biopotenciais podem ser mensurados (KANDEL *et al.*, 2000).

Dentre os sinais biopotenciais, pode-se mencionar os comumente utilizados em pesquisas na área da medicina: ECG que é um potencial elétrico proveniente de batidas cardíacas (JAMBUKIA *et al.*, 2015); EMG que se associa aos potenciais elétricos medidos no tecido muscular (YU *et al.*, 2019); EEG o qual mede atividades elétricas presentes no couro cabeludo (HOSSEINI *et al.*, 2021). Este último, é o sinal analisado na presente dissertação, portanto, faz-se necessário a fundamentação de suas principais características.

No sistema nervoso há a presença de células neuronais, que diferente de outras, tem uma estrutura complexa e polarizada. Os neurônios são compostos resumidamente por um corpo celular e por suas prolongações citoplasmáticas. O primeiro também é conhecido como soma e caracteriza a estrutura celular, já o segundo é dividido em duas categorias, dendritos e axônios (ELSEVIER, 2005). Os dendritos são ramificações presentes próximas ao corpo celular e desempenham papel de recebimento de impulsos nervosos e propagação para a soma (LONDON; HÄUSSER, 2005). Os axônios no que lhe concerne, tem o papel de transmitir o impulso proveniente do corpo celular até um próximo neurônio ou tecido, nota-se ainda que cada neurônio possui apenas um axônio (KIRKCALDIE; COLLINS, 2016).

A Figura 2 ilustra uma representação de um neurônio, em que é possível verificar que o mesmo é dividido em dois grandes grupos: somato-dendrítico (verde); axonal (azul). O

Figura 2 – Representação neuronal e seus principais grupos.



Fonte: Jones e Svitkina (2016). Modificado pelo autor.

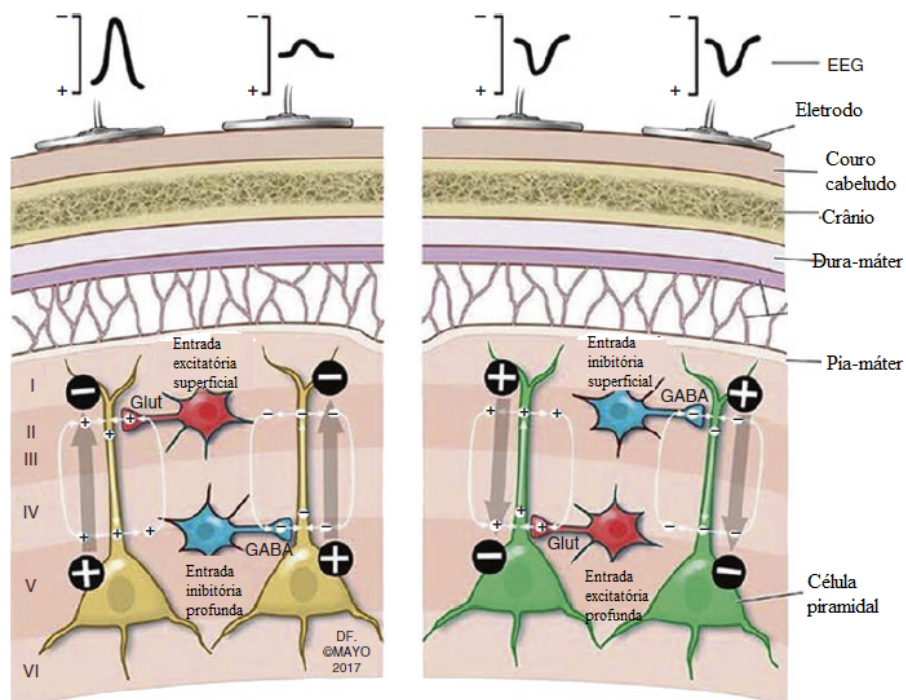
primeiro é composto pelo corpo celular, dendritos e o cone axonal. O segundo é referente a todos os elementos presentes no axônio. Essa distinção é realizada para caracterizar quando a atividade sináptica está sendo recebida ou transmitida. Portanto, os neurônios polarizados recebem o sinal pelo grupo somato-dendrítico que transmite para o segmento inicial do axônio através do cone axonal. Nesta etapa é iniciado um potencial de ação, que é propagado ao longo do axônio e o mesmo é amplificado ao passar pelos nódulos de Ranvier (JONES; SVITKINA, 2016).

As transmissões sinápticas entre neurônios podem ser de dois tipos, elétricas ou químicas. Para o primeiro caso, nos terminais pré e pós-sinápticos há um elemento de união que facilita a passagem da informação, portanto, sua transmissão ocorre em alta velocidade. Além disto, essas sinapses são comumente interconectadas em uma grande quantidade de neurônios multipolares. Por consequência, quando um limiar do potencial de ativação é atingido, o grupo de neurônios interconectados realiza o descarregamento de forma sincronizada (KANDEL *et al.*, 2000). Tratando das sinapses químicas, Kandel *et al.* (2000) comentam haver um espaçamento de magnitude de 30 – 50nm entre os terminais pré e pós-sinápticos. Neste caso, a atividade elétrica gerada pelo terminal axonal faz uma indução nas vesículas pré-sinápticas que produz um agente químico neuro-mediador no espaçamento, que tem papel facilitador para a comunicação com o terminal pós-sináptico.

Os neurônios piramidais encontrados no córtex cerebral possuem membranas pós-sinápticas que podem ser de caráter excitatório ou inibitório dependentes do neuro-mediador presente. Quando um potencial de ativação é atingido, a membrana inicialmente não permite novos estímulos e então gradualmente aumenta o limiar de excitabilidade. Porém, os neuro-mediadores que causam excitação podem fazer com que um *Excitatory PostSynaptic Potential*

(EPSP) seja realizado, aumentando assim a propensão em disparar um potencial de ação. De maneira contrária, se o neuro-mediador tem características inibitórias um *Inhibitory PostSynaptic Potential* (IPSP) pode ocorrer, diminuindo a possibilidade de potencial de ação. Os somatórios de EPSP e IPSP provenientes dos neurônios piramidais presentes abaixo do crânio, dá a resultante positiva e negativa respectivamente, dos registros EEG como pode ser visualizado através da Figura 3 (MECARELLI, 2019).

Figura 3 – Esquemático do procedimento de aquisição de EEG no couro cabeludo.



Fonte: Tatum *et al.* (2018). Modificado pelo autor.

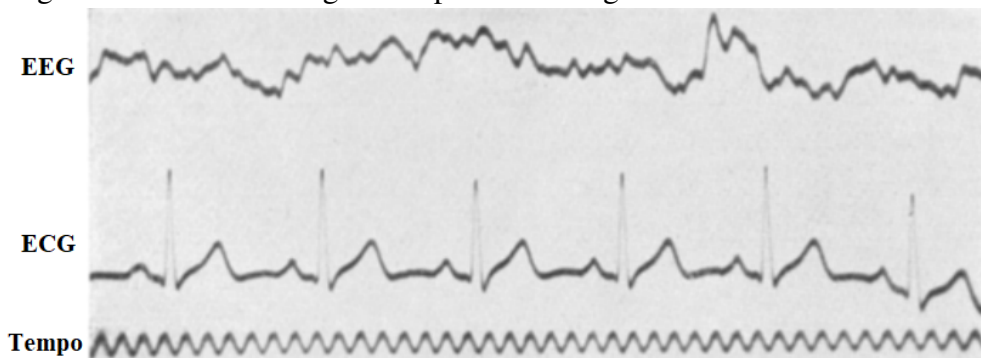
2.2 Método de Eletroencefalografia

A eletroencefalografia é um procedimento não invasivo que realiza medições dos potenciais elétricos cerebrais. De maneira geral, este método é realizado através de eletrodos acoplados na superfície do couro cabeludo, que são conectados a um dispositivo que realiza o registro das atividades provenientes do tecido neuronal (BIASIUCCI *et al.*, 2019). Porém, como será descrito, desde sua invenção, este método requer uma série de outros aparatos e protocolos para seu correto funcionamento.

O método foi criado por Hans Berger, que em 1924 realizou a primeira aquisição de

atividades cerebrais durante uma cirurgia neurológica. Nota-se que os procedimentos preliminares eram invasivos, portanto, os eletrodos eram posicionados diretamente no córtex cerebral. Porém, com avanços de seus estudos, Berger desenvolveu uma maneira de coletar os sinais EEG através do dispositivo galvanômetro de bonina dupla com dois eletrodos posicionados no couro cabeludo. Com isto, Berger foi capaz de expressar graficamente o sinal e realizar estudos sobre as características do mesmo. A Figura 4 ilustra um dos primeiros gráfico de sinal EEG registrado por Berger. Nesta é possível identificar que há a presença também do sinal ECG bem como a marcação do tempo de aquisição. Ressalta-se ainda que o sinal de EEG exibido pela Figura, é proveniente da derivação bipolar frontal-occipital (MECARELLI, 2019).

Figura 4 – Sinal EEG registrado por Hans Berger.



Fonte: Berger (1929). Modificado pelo autor.

Ao passar dos anos, pesquisas de modo a aprimorar esta técnica, foram desenvolvidas e atualmente um sistema que realiza as medições de EEG, possui diversas características como: expansão na quantidade de eletrodos e canais presentes; pré-amplificação e filtragem do sinal; conversor analógico-digital; *display* para exibição de informações (MECARELLI, 2019). Cada um destes subsistemas possui sua particularidade e podem ser utilizados de diferentes maneiras.

De maneira geral, os eletrodos são elementos que desempenham papel de conexão entre o paciente e o instrumento que grava os sinais bipotenciais, através da conversão do fluxo de íons presentes no corpo humano, em corrente elétrica. Isto ocorre através de reações químicas na interface eletrodo-eletrólito (NEUMAN *et al.*, 1998). De acordo com Webster (2009), as melhores interfaces para aquisição de sinais biopotenciais são aquelas combinadas de um metal e um de seus sais metálicos. Todavia, os eletrodos comumente possuem a composição de $Ag/AgCl$.

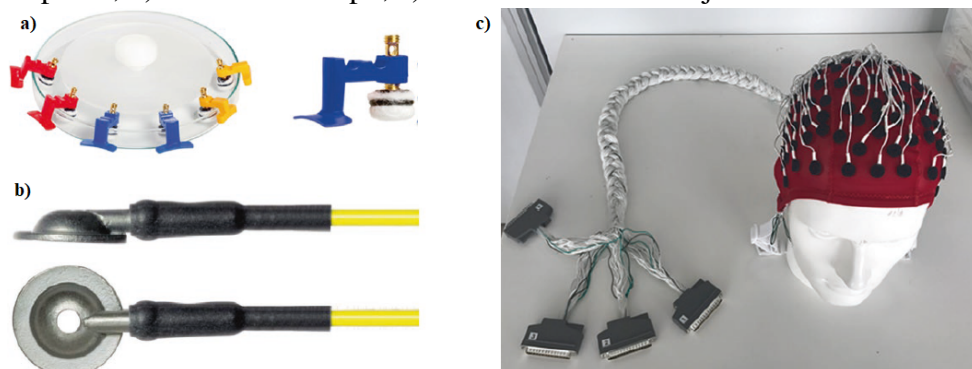
Eletrodos de ponte ou de disco são os mais comuns em aplicações que são realizadas aquisições de sinais EEG. O primeiro, ilustrado pela Figura 5a é composto normalmente de $Ag/AgCl$, possui um revestimento com algodão em volta de sua placa metálica e uma extensão

para acoplamento do cabo de comunicação. A maior vantagem do uso destes eletrodos está atrelada a flexibilidade no seu reposicionamento. Porém, ao se utilizar por um longo tempo o algodão resseca, aumentando assim sua impedância que causa artefatos (WEBSTER, 2009).

O eletrodo em formato de disco ilustrado pela Figura 5b, possui uma placa metálica circular e côncava com uma perfuração central. Nesta categoria de eletrodo, a extremidade oposta dos cabos conectados ao sistema de amplificação, são soldados diretamente no disco. A aplicação destes, no paciente pode ser realizada diretamente, ou em conjunto de uma gaze encharcada em colódio. Após isto, é inserido um gel condutor na perfuração central do disco. A principal diferença entre este eletrodo e o anterior é a sua não flexibilidade, porém, esta característica provê uma grande resistência aos movimentos do cabo, portanto, o uso destes eletrodos é recomendado para longos monitoramentos. Como principal desvantagem destaca-se o alto tempo de preparação para seu uso e a dependência do colódio, sendo uma solução adesiva tóxica e inflamável (WEBSTER, 2009).

Com os avanços do método de eletroencefalografia, os dispositivos se tornaram mais capazes de suportar diversas entradas e conseqüentemente mais eletrodos são utilizados para analisar diversos pontos do couro cabeludo simultaneamente. Desta maneira, foi desenvolvida uma touca ilustrada pela Figura 5c, a qual possui os eletrodos já fixados garantindo assim um procedimento mais confortável ao paciente além de um tempo de preparação menor (MECARELLI, 2019). Além disto, com esse aumento de eletrodos, foi necessária também uma padronização de posicionamento dos mesmos, de modo que pesquisadores pudessem comparar seus resultados com outros trabalhos (HOMAN, 1988).

Figura 5 – Eletrodos comumente utilizados na Aquisição EEG: a) Eletrodos de ponte; b) Eletrodo de copo; c) Touca com eletrodos já fixados.



Fonte: Mecarelli (2019). Modificado pelo autor.

Um dos padrões bastante utilizado é ilustrado através da Figura 6d. Este é chamado

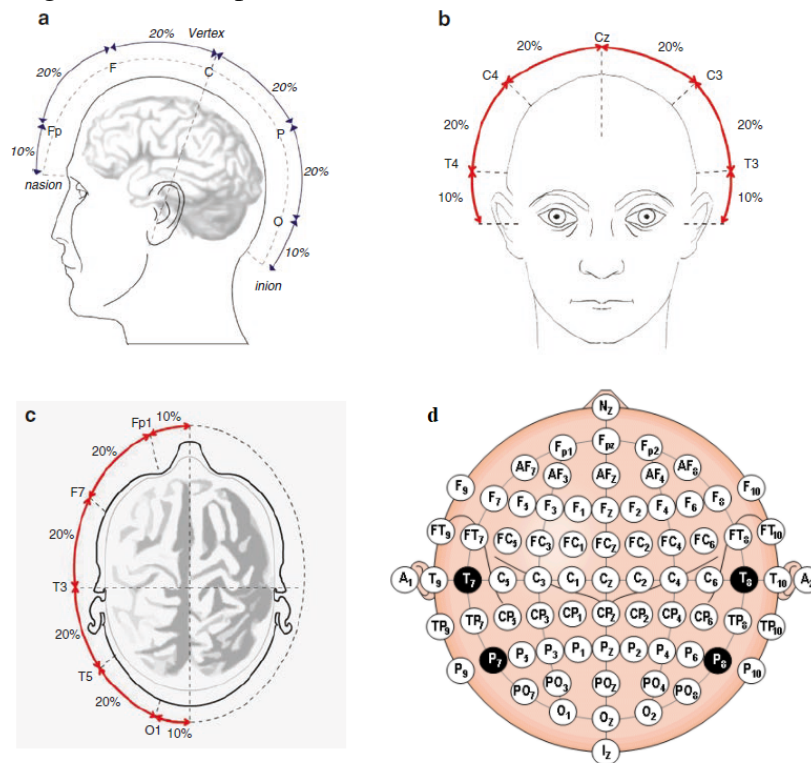
de sistema internacional 10 – 20 e foi proposto pela *International Federation of Societies for EEG and Clinical Neurophysiology* (IFSECN). De acordo com Jasper (1958), este padrão trata-se das definições de posicionamento dos eletrodos de uma maneira que os pontos sejam distribuídos ao longo do crânio proporcionalmente, de modo que se ocupe todo o couro cabeludo. Além disto em sua proposição, o sistema 10 – 20 faz a disposição de eletrodos intuitivamente de modo a destacar as principais áreas do cérebro com a letra O para occipital, F para frontal, C para central e P para parietal (JASPER, 1958).

De acordo com Jasper (1958), o posicionamento correto dos pontos propostos por esse padrão é realizado traçando linhas imaginárias, ligando pontos anatômicos no crânio. Na Figura 6a é possível verificar a primeira dessas linhas, traçada a partir da cavidade entre os olhos e acima do nariz até o osso que é ligeiramente projetado para fora do crânio no osso occipital. Nota-se que na linha, há cinco posições padronizadas e entre elas há um espaçamento de 20% do valor da distância total. Nas posições das extremidades até o ponto anatômico mais próximo, tem-se o espaçamento de 10% do total. A segunda linha é traçada perpendicularmente a primeira e tem como pontos anatômicos pré-auricular direito e esquerdo como pode ser visualizado pela Figura 6b. Além desta, pode-se traçar uma linha lateral longitudinal como ilustrado pela Figura 6c, que possui os mesmos pontos anatômicos descritos na primeira linha. Destaca-se que todos os espaçamentos de pontos, seguem o mesmo princípio explicado para a primeira linha, mantendo assim espaçamentos entre os posicionamentos com 20% e os espaçamentos entre posicionamento nas extremidades e pontos anatômicos com 10% da distância total da linha respectiva (KLEM, 1999).

Como dito anteriormente, os sistemas atuais são bastante desenvolvidos e possuem diversos aprimoramentos com relação aos iniciais. Uma das principais diferenças nestes sistemas, é o fato dos atuais serem digitais. Isto faz com que sejam bastante flexíveis ao modo de utilização da eletroencefalografia, principalmente na maneira de como gravar, analisar e exibir os dados. Um sistema dessa complexidade pode ser exemplificado através da Figura 7, que possui elementos como: sistema de aquisição de dados (1); acessórios para aquisição (2); dispositivos de processamento (3); equipamento que gera estímulos (4); periféricos de entrada (5) e saída (6) (MECARELLI, 2019).

Desses subsistemas pode-se destacar que o referente à aquisição, tem a função de receber os sinais analógicos ou imagens do paciente e convertê-los para o meio digital. O dispositivo 1 exibido na Figura 7 é comumente chamado de *headbox* e além de realizar o papel

Figura 6 – Exemplo do sistema internacional 10 – 20.



Fonte: Jasper (1958) e Shoeb (2009). Modificado pelo autor.

mencionado, possui circuitos internos que também realizam a amplificação do sinal bem como uma pré-filtragem do mesmo (MECARELLI, 2019). Geralmente esses subsistemas possuem grupos de amplificadores diferenciais normalmente operando com ganhos na magnitude de 10.000, tornando assim uma tensão que originalmente é na ordem de μV para uma tensão compatível com os circuitos digitais presentes. A motivação principal da utilização de uma pré-filtragem antes de enviar os sinais para a conversão analógico-digital é pelo simples fato que a etapa de amplificação também irá aumentar os ruídos (WEBSTER, 2009) (SHAD *et al.*, 2020). A principal causa destes ruídos está relacionada ao contato do eletrodo com o couro cabeludo, que devido a sua impedância pode causar artefatos e indução eletromagnética para os cabos, afetando assim os sinais. Sabe-se que os sinais EEG tem uma amplitude baixa, na ordem de μV , portanto, o mesmo é bastante suscetível a este a ruídos desta natureza. As induções que causam ruídos numa frequência de $60Hz$, são provenientes da rede elétrica e portando os sistemas normalmente possuem um filtro rejeita faixa para eliminar esses sinais (PISE; REGE, 2021).

Figura 7 – Sistema completo para eletroencefalografia.



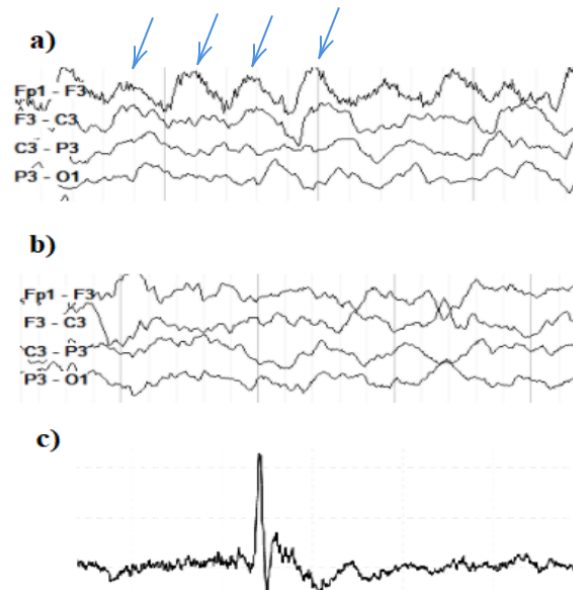
Fonte: (MECARELLI, 2019). Modificado pelo autor.

2.3 Características de sinais EEG

Com o procedimento da aquisição dos sinais EEG já descrito, apresentam-se agora as principais características do sinal. Normalmente um paciente sob exame de eletroencefalografia, deve se mantêm em estado de repouso ou realizando alguma tarefa específica. Desta maneira, o sistema consegue analisar diversos aspectos sobre os sinais EEG. Tais características são, sua forma de onda, amplitude, distribuição de frequência, distribuição espacial e simetria entre hemisférios cerebrais. Todas essas características são variantes e dependem da idade, estado mental, nível de consciência, condições fisiológicas e características individuais da pessoa (SAVADKOOHI *et al.*, 2020).

De acordo com Mecarelli (2019), as formas de onda podem ser classificadas como lenta ou aguda. A primeira representa sinais que possuem uma longa duração na casa dos milissegundos e podem ser categorizadas como monomórficas para sinais que possuem uma certa periodicidade, ou caso contrário são ditas polimórficas, estas são ilustradas pelas Figuras 8a

Figura 8 – a) Sinal lento monomórfico; b) Sinal lento polimórfico; c) Sinal agudo.



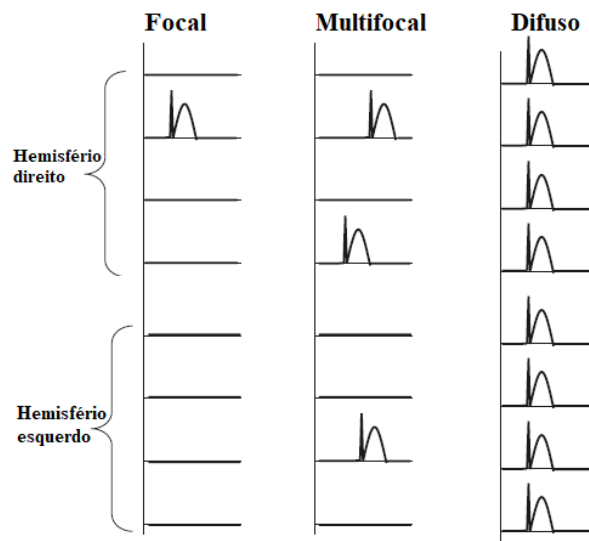
Fonte: (VALENTINE, 2020) e (GOLOVKO; BEZOBRANOVA, 2021). Modificado pelo autor.

e b respectivamente. Os sinais classificados como agudos tem uma duração curta como ilustrado pela Figura 8c (KHOUMA *et al.*, 2019).

Com relação às análises extraídas da distribuição espacial, pode-se dizer que algumas vezes um sinal EEG é observado simultaneamente em diferentes locais dos hemisférios, de modo síncrono. Todavia, esses sinais também podem ocorrer de modo assíncrono quando aparecem em locais diferentes com certo atraso temporal (TABERNA *et al.*, 2019). Com isto dito, é importante na análise espacial, conhecer a localização dos eletrodos e sua derivação no couro cabeludo. Logo, neste tipo de análise é possível caracterizar os sinais como ilustrado pela Figura 9. Portanto, o sinal pode ser definido como focal quando é presente em uma única área de um hemisfério cerebral, multifocal quando é observado em diferentes áreas dos hemisférios de modo assíncrono e por último podem ser categorizados como difuso quando são presentes de modo síncrono nas áreas de ambos os hemisférios (SEECK *et al.*, 2017).

Uma das características mais importantes na análise dos sinais de EEG são suas distribuições de frequência, que estão agrupadas em faixas como pode ser visualizado no Tabela 1. Como dito anteriormente, um procedimento comumente realizado é o pré-processamento do sinal, removendo as componentes cujas faixas de frequência estão acima de 30Hz e abaixo de $0,5\text{Hz}$ através de filtro passa-faixas. Porém, como será descrito, em alguns casos esse limite inferior pode ser estendido para abranger frequências menores (MECARELLI, 2019; ACHARYA

Figura 9 – Sistema completo para eletroencefalografia.



Fonte: Mecarelli (2019). Modificado pelo autor.

Tabela 1 – Faixas de Frequência de sinais EEG.

	Limite Inferior (Hz)	Limite Superior (Hz)
<i>delta</i> (δ)	0,1	< 4
<i>teta</i> (θ)	4	< 8
<i>alfa</i> (α)	8	< 13
<i>beta</i> (β)	13	30

Fonte: Mecarelli (2019).

et al., 2015).

As ondas *delta* são as mais lentas e possuem menor faixa de frequência, como pode ser visualizado através da Figura 10a. Por conta de sua morfologia, ao realizar inspeção visual do sinal, enquanto o paciente está acordado, muitas vezes as componentes nessa faixa de frequência são mascaradas pelas demais. Estas ondas podem ser verificadas com clareza quando há movimentação de olhos enquanto o paciente está em estado profundo de sono (TONONI; CIRELLI, 2012). Como dito anteriormente, algumas análises são realizadas ao aumentar a gama da faixa de frequências baixas, ao analisar também as inferiores a $0,5Hz$. Porém, este caso não tem um papel ativo em aplicações clínicas, mas sim pesquisas como, por exemplo, o estudo da atividade cortical ou identificação da zona epiletogênica (HENRY, 2006).

Um exemplo de ondas *teta* (4 até $< 8Hz$) pode ser visualizado através da Figura 10b. Estas são dependentes das características individuais ou nível de vigiância do sujeito em análise. Estas faixas de frequências são mais representadas em crianças do que em adultos, já que nestes

elas são um indicativo de anormalidade (FATTINGER *et al.*, 2017). Porém, são verificadas quando um paciente está realizando tarefas para resolver problemas ou quando há ocorrência de ansiedade, ou dor (NIEDERMEYER; SILVA, 2020). Além disto, são visualizadas quando um paciente que outrora estava acordado entra num estado de sonolência (VYAZOVSKIY; TOBLER, 2005).

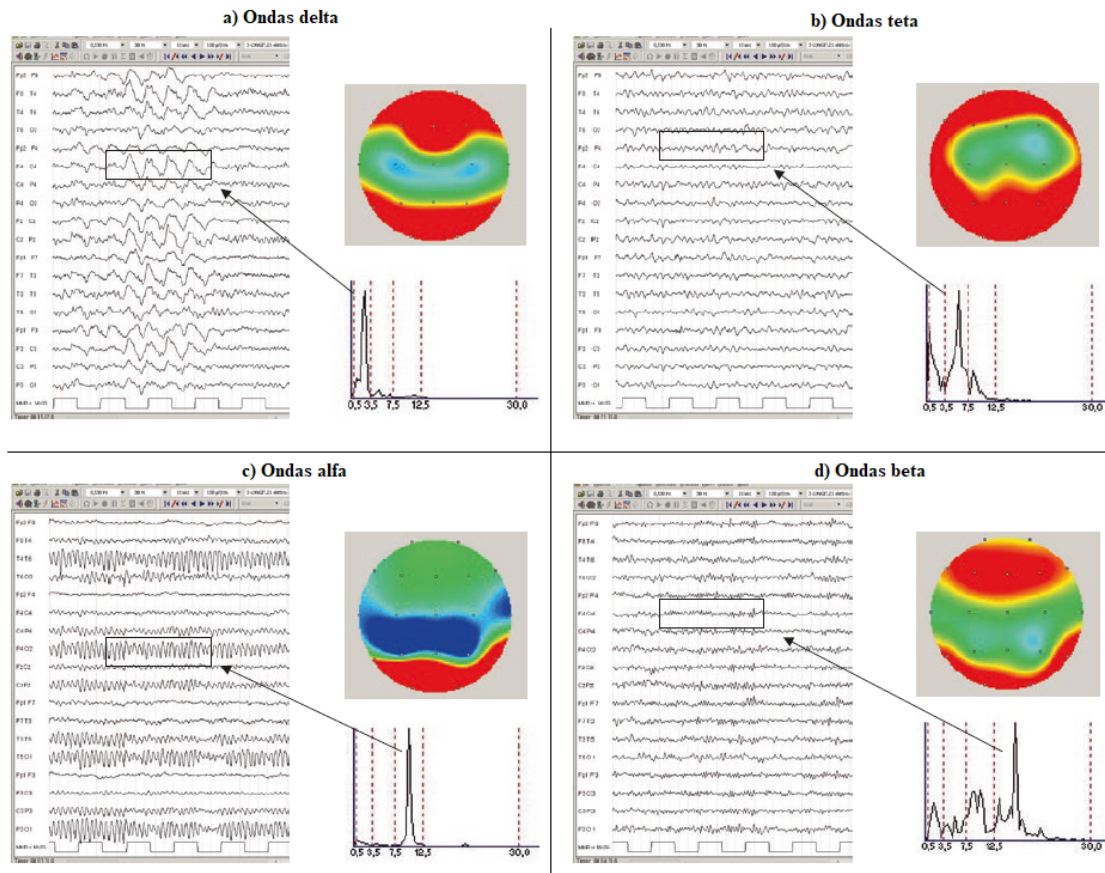
As ondas *alfa* (8 até $< 13\text{Hz}$) que também foram inicialmente descritas por Berger, são visualmente percebidas em crianças e adultos acordados e com olhos fechados (KANE *et al.*, 2017). Estas ondas são verificadas comumente na região occipital, porém, há casos onde esse ritmo pode se espalhar até as áreas frontal e central. A morfologia destas ondas tem característica monomórfica e senoidal como pode ser visualizado na Figura 10c (MECARELLI, 2019). Estas ondas são extensivamente analisadas em ambientes clínicos, já que possuem variações baseadas no estado cognitivo do paciente e por estarem presentes durante maior parte da vida humana. Além disto, essas características mencionadas podem ser visualizadas também em pacientes cegos de nascença (BAYS *et al.*, 2015).

Por último, a frequência *beta* (13 até 30Hz) é identificada nas regiões frontal e central temporal de adultos acordados em repouso. Estas ondas podem ser divididas em duas faixas, a primeira com intervalos de 13 a 21Hz associa-se a atividade beta de paciente saudável, já a segunda que varia entre 22 a 30Hz está relacionada à ansiedade, estresse ou excesso de estímulos (M. *et al.*, 2019). Portanto, num geral as ondas betas estão relacionadas a certos processos cognitivos ou mentais, porém sua atividade pode ser distorcida por outros fatores como, interferência gerada pela atividade muscular ou até mesmo na utilização de algum medicamento ou drogas (NIEDERMEYER; SILVA, 2020).

Como dito anteriormente, os sinais EEG operam em baixa frequência ($< 30\text{Hz}$), portanto, são bastantes suscetíveis a ruídos. Estes sinais indesejáveis são chamados de artefatos, e podem ser caracterizados como sinais de origem não cerebrais (JAFARIFARMAND; BADAM-CHIZADEH, 2013). Comumente os artefatos são classificados em três categorias que os causam: processos fisiológicos; pelo sistema de aquisição; por interferência elétrica ou equipamentos externos (URIGÜEN; GARCIA-ZAPIRAIN, 2015).

Os artefatos fisiológicos são uma categoria de sinais ruidosos provenientes de processos diferentes dos cerebrais. Destaca-se que os causadores de maior interferência, são provenientes da movimentação de olhos, atividades musculares, batimentos cardíacos e respiração (URIGÜEN; GARCIA-ZAPIRAIN, 2015; KAPPEL *et al.*, 2017). Movimentos oculares são

Figura 10 – Exemplos de sinais por faixa de frequência.



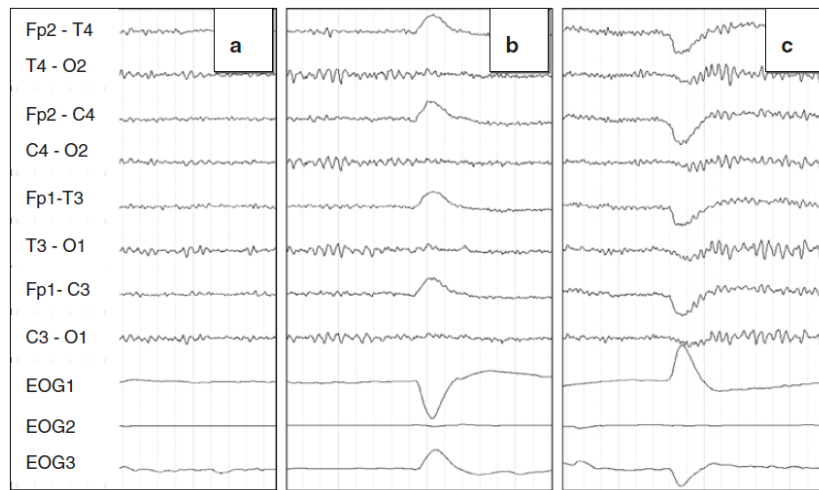
Fonte: Mecarelli (2019). Modificado pelo autor.

geralmente captados pelos eletrodos posicionados no lobo frontal, porém, podem ser percebidos em outras regiões. A maior incidência dessa interferência é causada na piscada de olhos como pode ser visualizado na Figura 11. Nesta Figura é possível verificar uma comparação entre os sinais de EEG e Eletrooculograma (EOG). Portanto, ao abrir ou fechar os olhos percebe-se principalmente pelos eletrodos Fpx, uma característica de sinal agudo fora do padrão no sinal EEG (MECARELLI, 2019).

Os artefatos causados pela atividade muscular, tem origem principalmente nos grupos musculares presentes na região temporal, na face e no pescoço. O sinal desses artefatos tem como características, uma alta frequência e amplitudes superiores ao sinal EEG, contudo, não possuem aspectos morfológicos repetitivos (BRIENZA *et al.*, 2019). Uma abordagem para evitar esses artefatos, é realizar a aquisição dos sinais quando o paciente está relaxado, com a boca levemente aberta e em uma posição que garanta com que seu pescoço não esteja contraído (NIEDERMEYER; SILVA, 2020).

Por último, os sinais de ECG podem também contaminar os sinais de EEG. Este tipo

Figura 11 – Artefatos por movimento ocular.



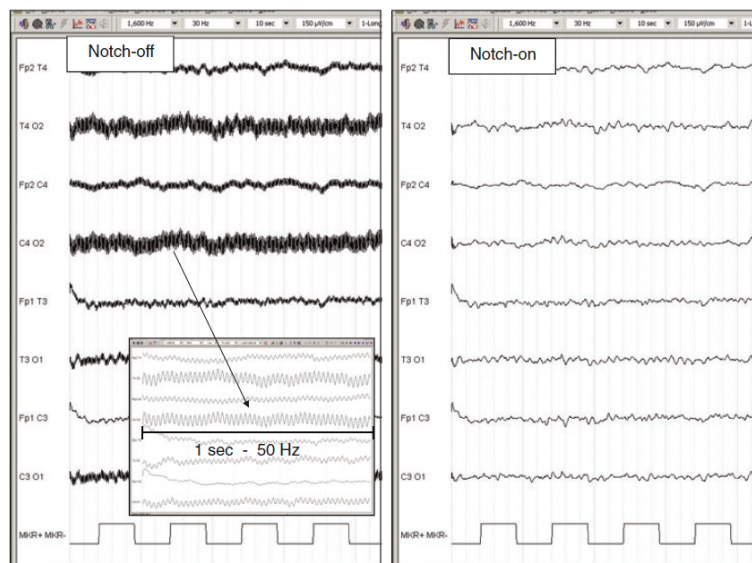
Fonte: Mecarelli (2019). Modificado pelo autor.

de artefato é produzido pelo campo elétrico gerado pela atividade cardíaca sendo considerado um dos grandes causadores de diagnósticos imprecisos, pela dificuldade em eliminá-los dos sinais EEG. Portanto, normalmente a abordagem é monitorar as atividades cardíacas simultaneamente com as cerebrais (DORA; BISWAL, 2016; NIEDERMEYER; SILVA, 2020).

Ao se tratar da categoria de artefatos produzidos pelo sistema de aquisição, comumente o motivo da presença de ruídos está relacionada com os eletrodos. Pode-se destacar que existem diversos motivos para que um eletrodo não desempenhe corretamente suas funções, porém, os principais que causam os artefatos são o uso excessivo de material condutor e mau posicionamento de eletrodos. Além disto, outro artefato desta categoria é produzido por oscilações nos cabos do sistema. Esta no que lhe concerne, pode ser ocasionada pela movimentação do paciente, bem como por algum tremor ou fluxo de ar na sala (BRIENZA *et al.*, 2019).

Por fim, a última categoria enquadra os artefatos produzidos por outros equipamentos ou interferências elétricas. O ruído mais comum neste grupo é causado pela rede elétrica, já que suas componentes têm frequências superiores (50 ou 60Hz) dos sinais de EEG (< 30Hz). Um exemplo da contaminação através deste ruído, pode ser visualizada através da Figura 12. Nesta, é possível verificar a diferença do sinal quando o filtro rejeita-banda é aplicado, removendo assim os ruídos causados pelos 50Hz. Como normalmente as aquisições dos sinais EEG são realizadas em ambiente hospitalar, esses artefatos podem ser captados pela operação de outros dispositivos (BRIENZA *et al.*, 2019). Pode-se destacar que o uso de telefone fixo, ou o ligar e desligar de equipamentos (até mesmo lâmpadas) também pode corromper os sinais EEG (MECARELLI, 2019).

Figura 12 – Artefato causado pela interferência da rede elétrica.



Fonte: Mecarelli (2019). Modificado pelo autor.

2.4 Considerações Finais

No presente capítulo, foram descritos os fundamentos relacionados aos sinais de EEG, expuseram-se como estes sinais são captados, quais dispositivos mais comumente utilizados, quais os padrões utilizados para aquisição do sinal, e por fim, foi descrito sobre as morfologias do sinal EEG e suas problemáticas. O próximo capítulo no que lhe concerne, traz informações sobre os métodos clássicos para extrair informações úteis do sinal EEG. Descreve-se nesse capítulo sobre dois métodos amplamente utilizados e como estes são aplicados em uma massa de dados de EEG.

3 EXTRAÇÃO DE ATRIBUTOS DE SINAIS DE EEG: MÉTODOS CLÁSSICOS

Por conta das características do sinal EEG descritas no capítulo 2, há diversas dificuldades enfrentadas por profissionais da saúde no momento da análise destes sinais. Pode-se destacar que a tomada de decisões através de uma inspeção visual, além de ser demorada, é difícil de ser realizada. Com o início de pesquisas utilizando algoritmos de AM, os pesquisadores conseguiram propor diversas técnicas na tentativa de potencializar o desempenho de modelos, que resolvem problemas relacionados ao reconhecimento de padrões (REN; HAN, 2018). Como os desempenhos desses algoritmos na etapa de classificação são dependentes da natureza dos dados a serem analisados, faz-se o uso de técnicas para extrair as suas informações relevantes (SAIRAMYA *et al.*, 2018).

Estas técnicas são chamadas de extração de atributos e para os sinais EEG, podem ser chamadas parametrização. Com esta é possível realizar uma transformação do sinal original em vetores de atributos rotulados. Estes por sua vez, são agrupados e podem ser aplicados nos algoritmos de AM. Comumente os métodos de extração de atributos são categorizados em lineares ou não lineares. Os primeiros são baseados em modelos lineares e analisam os sinais de EEG nos domínios do tempo, frequência ou em ambos. Os métodos não lineares por sua vez, analisam as propriedades dinâmicas não lineares dos sinais de EEG (REN; HAN, 2018).

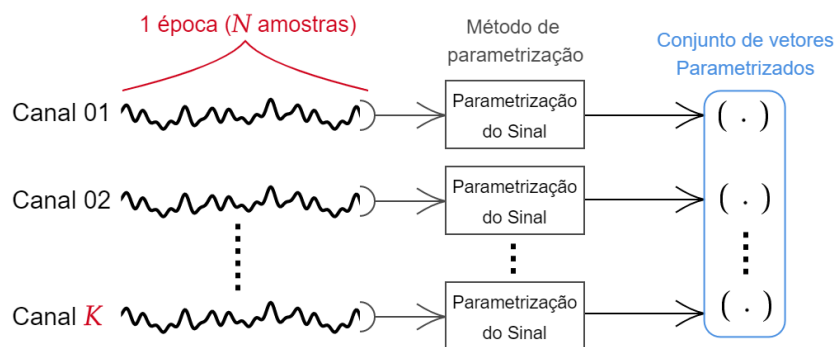
Com isto dito, apresentam-se dois métodos clássicos amplamente utilizados para parametrizar os sinais de EEG. Esta fundamentação é necessária, pois a hipótese inicial do presente trabalho busca uma parametrização superior ou tão boa quanto às providas por estes métodos clássicos. Inicialmente, faz-se necessário algumas definições para padronizar certas nomenclaturas ao longo da presente dissertação, bem como as abordagens de aplicação destes métodos clássicos. Estes por sua vez, são descritos em sequência, dos quais o primeiro utiliza informação temporal explícita de modo intracanal, ao passo que o segundo em sua primordialidade captura a informação espacial pela abordagem intercanais. Todavia, este último no que lhe concerne, também pode ser aplicado de modo intracanal como será descrito.

Os sinais de EEG são capturados através de um sistema que normalmente possui uma quantidade K de canais. Ao longo do tempo, a quantidade de sinais adquiridos está em função da taxa de amostragem do equipamento. Normalmente o valor em Hz desta taxa é baixo ($SR \leq 512Hz$), de modo que a cada segundo provê uma aquisição de SR amostras. A análise dos sinais de EEG é realizada em curtos segmentos do sinal (1 ou 2s) chamadas épocas. Portanto, considerando um exemplo prático, a quantidade de amostras em uma época de t segundos é

$N = t \cdot SR$ em cada canal.

A parametrização do sinal é realizada a cada época e sua aplicação pode ser de modo intracanal ou intercanais. A primeira abordagem, é ilustrada através da Figura 13 que exhibe o processo de parametrização em apenas uma época. É possível verificar que neste modo, o método que parametriza o sinal é aplicado em cada canal individualmente e com isto tem-se K vetores parametrizados. Quando o sinal referente ao último canal é parametrizado, todos os vetores são concatenados verticalmente e assim este pode ser rotulado.

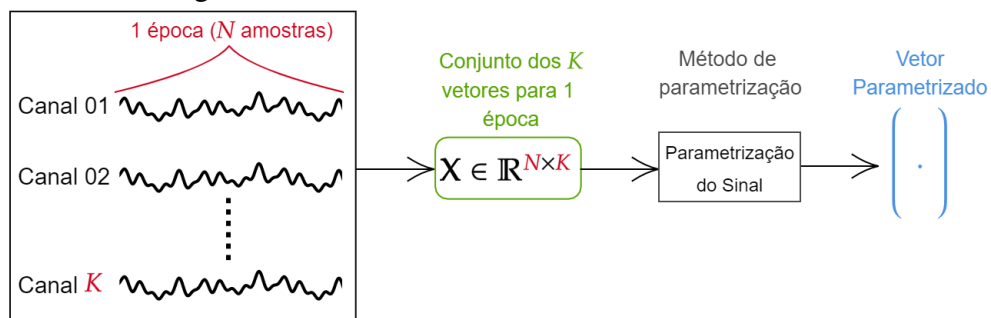
Figura 13 – Abordagem intracanal.



Fonte: Elaborado pelo autor.

A abordagem pelo modo intercanais é ilustrada através da Figura 14. Nesta, é possível verificar que todas as N amostras em uma época, para um dado canal, são organizadas como uma coluna da matriz $\mathbf{X} \in \mathbb{R}^{N \times K}$. Esta por sua vez, é enviada ao método de parametrização, que produz um único vetor com as informações relevantes, e pode ser rotulado numa próxima etapa.

Figura 14 – Abordagem intercanais.



Fonte: Elaborado pelo autor.

Com estas definições estabelecidas, apresentam-se os métodos LPC que se utiliza da

abordagem intracanal e o de covariância que pode ser aplicado em ambos modos, intracanal ou intercanais.

3.1 Vetor de Coeficientes LPC

A extração de atributos através dos coeficientes LPC, faz uma análise no domínio do tempo e é baseado no método Autorregressivo (AR). Tal método é amplamente utilizado em ajustes de séries temporais e pode descrever bem os sinais de EEG, já que para análises em curtos intervalos de tempo, pode-se assumir a estacionariedade. O método AR pode ser descrito como uma função que modela p valores anteriores de uma série temporal $\mathbf{x}_t = \{x(1), x(2), \dots, x(N)\}$ (SHUMWAY, 2017). Esta técnica linear é conhecida como coeficientes LPC sendo aplicada através da abordagem intracanal. O filtro AR pode ser definido conforme a seguinte equação:

$$x(t) = a_1x(t-1) + a_2x(t-2) + \dots + a_px(t-p) + v(t), \quad (3.1)$$

em que p representa a ordem do filtro, a_1, a_2, \dots, a_p são os coeficientes LPC e $v(n)$ equivale um processo de ruído branco gaussiano (REN; HAN, 2018).

Para utilizar o método LPC a fim de parametrizar o sinal de EEG, é necessário inicialmente a escolha da ordem do filtro p , e em seguida realizar a estimação dos respectivos coeficientes LPC. Esta estimação pode ser realizada de diferentes formas, porém, os métodos comumente utilizados são: Burg; Yule-Walker; Mínimos Quadrados (ZAZZARO *et al.*, 2021);(SHUMWAY, 2017). Para fins do presente trabalho, descreve-se a estimação através do método dos Mínimos Quadrados.

A estimação dos coeficientes LPC através do método mencionado, faz a minimização da soma do erro quadrático de predição. Com o valor de p definido, faz-se a composição de uma matriz de Toeplitz e um vetor de predições, a partir de uma série temporal. Portanto, ao considerar um segmento de sinal EEG proveniente de apenas um canal, ou seja $\mathbf{x}_i = \{x(1), x(2), \dots, x(N)\} \in \mathbb{R}^{1 \times N} | i = \{1, \dots, K\}$, pode-se montar o vetor de predições $\mathbf{p} \in \mathbb{R}^{(N-p) \times 1}$, de acordo com

$$\mathbf{p} = \begin{bmatrix} x(p+1) \\ x(p+2) \\ \vdots \\ x(N) \end{bmatrix}, \quad (3.2)$$

e a matriz de Toeplitz $\mathbf{T} \in \mathbb{R}^{(N-p) \times p}$ por sua vez, pode ser composta de acordo com

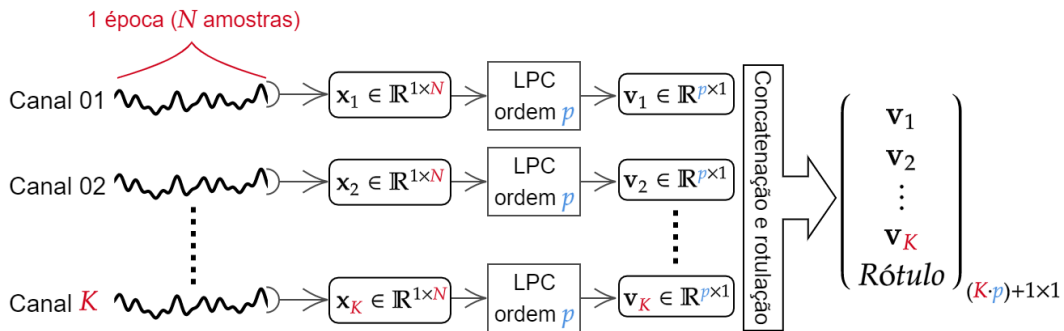
$$\mathbf{T} = \begin{bmatrix} x(p) & x(p-1) & \cdots & x(1) \\ x(p+1) & x(p) & \cdots & x(2) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-1) & x(N-2) & \cdots & x(N-p) \end{bmatrix}. \quad (3.3)$$

Por fim, o vetor $\hat{\mathbf{v}} \in \mathbb{R}^{p \times 1}$ que contém os coeficientes LPC do canal analisado pode ser estimado por meio da equação (BÖTTCHER, 1999; SHUMWAY, 2017):

$$\hat{\mathbf{v}} = (\mathbf{T}^T \mathbf{T})^{-1} \mathbf{T}^T \mathbf{p}. \quad (3.4)$$

O procedimento para obtenção do vetor final parametrizado referente aos K canais em uma determinada época, pode ser visualizado através do fluxograma ilustrado pela Figura 15. Verifica-se através desta, que após a aplicação do método em cada canal, faz-se uma concatenação dos vetores resultantes, formando um único vetor de ordem $\mathbb{R}^{(K \cdot p) \times 1}$. Este no que lhe concerne, caracteriza o fim da época em questão e, portanto, deve ser rotulado formando assim o vetor final com ordem $\mathbb{R}^{(K \cdot p) + 1 \times 1}$. O procedimento então é repetido para as demais épocas concatenando os vetores de atributos rotulados pela coluna.

Figura 15 – Fluxograma da parametrização via LPC de uma época.

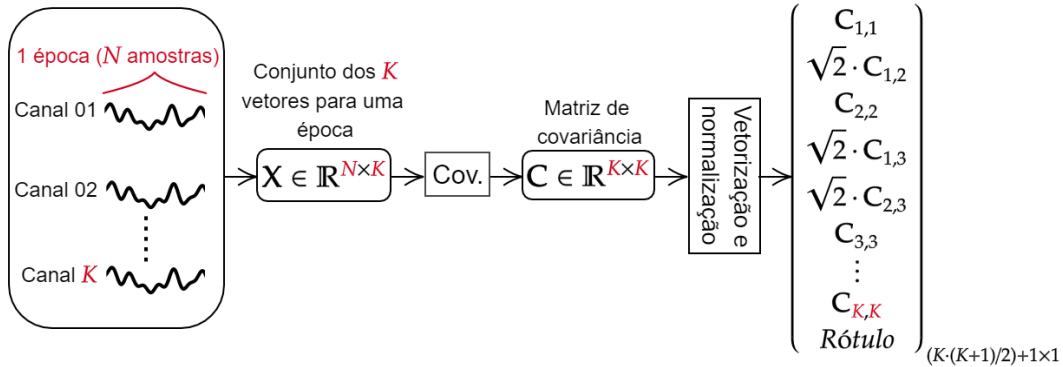


Fonte: Elaborado pelo autor.

3.2 Vetor de Covariâncias

Outra abordagem clássica para parametrização de sinais EEG pode ser realizada através de uma análise espacial e estatística. Ou seja, todos os canais são avaliados simultaneamente para determinada época (MILLIKEN, 1984; BARACHANT *et al.*, 2013). Tal procedimento pode ser visualizado através da Figura 16, que ilustra o fluxo para obtenção do vetor de características

Figura 16 – Fluxograma da parametrização via covariância pelo modo intercanais.



Fonte: Elaborado pelo autor.

rotulado. É possível verificar que inicialmente usa-se a massa de dados $\mathbf{X} \in \mathbb{R}^{N \times K}$ referentes à uma determinada época para estimar a matriz de covariância $\mathbf{C} \in \mathbb{R}^{K \times K}$. Em sequência, esta é normalizada e vetorizada de maneira que seja produzido o vetor parametrizado que pode ser rotulado.

A estimação da matriz de covariância pode ser realizada através da seguinte equação:

$$\mathbf{C} = \frac{1}{N} \sum_{t=1}^N [\mathbf{x}_t - \mathbf{m}]^T [\mathbf{x}_t - \mathbf{m}], \quad (3.5)$$

que utiliza da massa de dados \mathbf{X} e do vetor $\mathbf{m} \in \mathbb{R}^{1 \times K}$. Este último também necessita ser estimado e representa o vetor médio das N amostras em cada canal K (NIK, 2014). A matriz de covariância tem como características, ser simétrica e possuir informações de variância entre canais em sua diagonal principal e covariância de pares nos demais elementos. Essa é uma propriedade interessante, pois a matriz de covariância permanece inalterada se por acaso os vetores em \mathbf{X} fossem embaralhados (KOLKHORST *et al.*, 2020).

Como se trata de uma abordagem para extração de atributos, é interessante realizar um procedimento de vetorização da matriz \mathbf{C} de modo a produzir o vetor de características referente a uma época. Como dito anteriormente, a matriz de covariância é simétrica, portanto, faz-se necessário o procedimento de meia-vetorização. Este faz a escolha dos valores presentes na matriz triangular superior ou inferior, para compor os dados no vetor parametrizado. Tal composição forma o vetor $\mathbf{v} \in \mathbb{R}^{[(K+1)K/2] \times 1}$ já ilustrado pela Figura 16 sendo realizada através da seguinte equação:

$$\mathbf{v} = \left[\hat{C}_{1,1} \quad \sqrt{2} \cdot \hat{C}_{1,2} \quad \hat{C}_{2,2} \quad \sqrt{2} \cdot \hat{C}_{1,3} \quad \sqrt{2} \cdot \hat{C}_{2,3} \quad \hat{C}_{3,3} \quad \cdots \quad \hat{C}_{K,K} \right]^T. \quad (3.6)$$

Nesta é possível verificar que os elementos são multiplicados $\sqrt{2}$ com exceção daqueles presentes na diagonal principal. Isto é realizado de modo a preservar a igualdade entre as normas $\|\hat{\mathbf{C}}\|_F = \|\mathbf{v}\|_2$. Na qual $\|\cdot\|_F$ denomina a norma de Frobenius e $\|\cdot\|_2$ representa a norma euclidiana (BARACHANT *et al.*, 2013).

Como fora mencionado, há uma possibilidade de aplicação do método de covariância pelo modo intracanal, a fim de parametrizar o sinal EEG individualmente em cada canal. Para isto, é necessário que cada vetor $\mathbf{x} \in \mathbb{R}^{1 \times N}$ seja transformado em uma matriz de regressores (ou Toeplitz) $\mathbf{R} \in \mathbb{R}^{(N-p+1) \times p}$, em que p é um parâmetro à ser definido. Destaca-se que o procedimento para montagem da matriz é o mesmo utilizado no método LPC, porém, com a diferença de possuir uma linha a mais, como pode ser verificado através da seguinte equação:

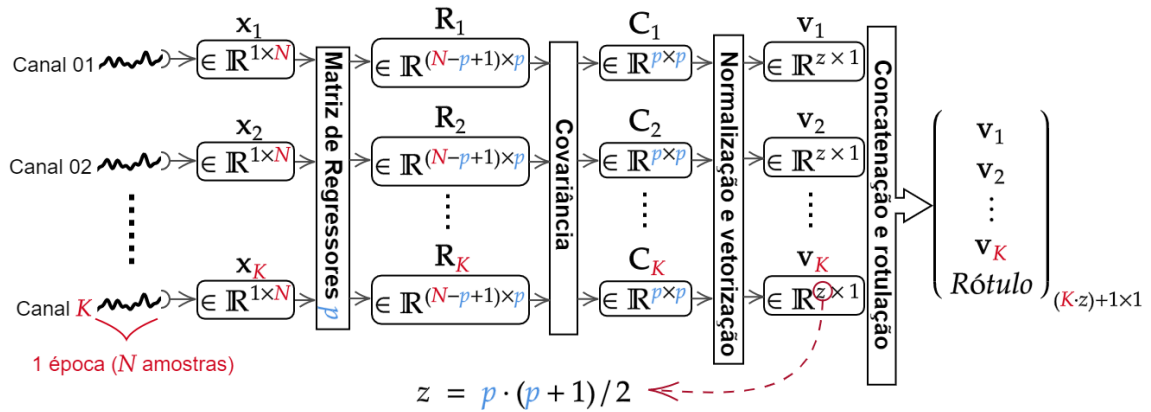
$$\mathbf{R} = \begin{bmatrix} x(p) & x(p-1) & \cdots & x(1) \\ x(p+1) & x(p) & \cdots & x(2) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-1) & x(N-2) & \cdots & x(N-p) \\ x(N) & x(N-1) & \cdots & x(N-p+1) \end{bmatrix}. \quad (3.7)$$

Em sequência, faz-se a estimação da matriz de covariância, que para este caso, é $\mathbf{C} \in \mathbb{R}^{p \times p}$. Esta por sua vez é normalizada e vetorizada da mesma maneira como é feito para o modo intercanais, produzindo assim o vetor parametrizado $\mathbf{v} \in \mathbf{R}^{(p \cdot (p+1)/2) \times 1}$ referente a apenas um canal. Logo, repete-se este procedimento para todos os canais formando K vetores parametrizados. Por fim, como pode ser visualizado através da Figura 17, faz-se a concatenação dos K vetores em um único de ordem $[K \cdot (p \cdot (p+1)/2)] + 1 \times 1$ que é rotulado com sua classificação. Da mesma maneira como é realizado no método LPC, o procedimento descrito é repetido para todas as épocas, concatenando os vetores de atributos rotulados pela coluna.

3.3 Considerações Finais

O presente capítulo abordou sobre a extração de atributos em sinais de EEG. Foram escolhidos dois métodos clássicos amplamente utilizados para parametrizar tais sinais. O primeiro, chamado vetor de coeficientes LPC, trata-se de um método linear baseado numa análise temporal intracanal. O segundo método que chamado vetor de covariâncias, no que lhe concerne, trata-se de uma análise espacial estatística e pode ser aplicado em modo intracanal ou intercanais. O próximo capítulo é caracterizado pelos métodos de parametrização de sinais de EEG propostos pela presente dissertação. Diferente dos métodos clássicos, os apresentados no capítulo que se

Figura 17 – Fluxograma da parametrização via covariância pelo modo intracanal.



Fonte: Elaborado pelo autor.

segue, são não lineares, porém, ainda fazem o uso das abordagens intercanais e intracanal. Há ainda ao final do capítulo uma seção em que se descreve sobre os classificadores utilizados no presente trabalho para avaliar os desempenhos dos métodos de parametrização.

4 EXTRAÇÃO DE ATRIBUTOS DE SINAIS DE EEG: MÉTODOS PROPOSTOS

Como dito anteriormente, o presente trabalho apresenta como proposta uma abordagem nova de extração de atributos de sinais EEG. Esta abordagem de parametrização é realizada através do uso de matrizes de kernel em conjunto de métodos de esparsificação. O presente trabalho tem como tentativa a verificação de que estes métodos, que são não lineares, sejam tão bons ou melhores do que os clássicos e lineares descritos no Capítulo 3.

Inicialmente, são abordados os conceitos em torno das matrizes de kernel. Nesta etapa, descrevem-se os fundamentos para estimação destas matrizes através de duas funções de kernel amplamente utilizadas. Em sequência explica-se sobre os métodos de esparsificação utilizados de modo a reduzir a alta dimensionalidade das matrizes de kernel ao produzir um dicionário de vetores relevantes.

4.1 Extração através de matrizes de kernel

Para descrever como as matrizes de kernel podem ser utilizadas com o propósito de parametrização do sinal EEG, faz-se necessário inicialmente a fundamentação do seu conceito. Em sequência é explicado como é realizada sua estimação, e por fim quais são as abordagens comuns ao se utilizar essas matrizes. De acordo com Svensén e Bishop (2007) o conceito kernel foi introduzido ao tema de reconhecimento de padrões por volta de 1964 em pesquisas com eletrostática através de método de função potencial. Porém, o uso dessas matrizes foi amplamente difundido no início da década de 90 no AM em pesquisas (como a de BOSER *et al.*, 1992) de classificadores baseados em maximização de margem, que deram origem ao método *support vector machines* (SVM).

Pode-se destacar que um dos problemas fundamentais do reconhecimento de padrões, é fazer com que um modelo gerado através de um conjunto conhecido, consiga fazer predições ou generalizações similares aos exemplos conhecidos. Portanto, nesse contexto, faz-se necessária a definição da medida de similaridade. Uma das maneiras mais simples de se obter esta medida é através do produto interno entre vetores. Ou seja, dados dois vetores \mathbf{x} e \mathbf{x}' de mesma ordem, o produto interno representado pela seguinte equação:

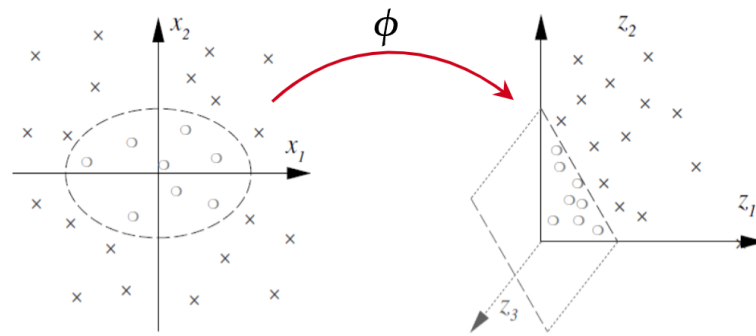
$$\langle \mathbf{x}, \mathbf{x}' \rangle = \sum_{i=1}^N x_i \cdot x'_i = \mathbf{x}^T \mathbf{x}', \quad (4.1)$$

sendo uma função que tem como retorno um valor escalar que caracteriza a similaridade entre os

vetores, e tem como interpretação geométrica, o produto das magnitudes dos vetores \mathbf{x}, \mathbf{x}' e o cosseno do ângulo entre eles (SCHÖLKOPF *et al.*, 2002).

A partir desta explicação de similaridade, é possível descrever que o conceito kernel é baseado na utilização desta métrica, porém através de uma abordagem dual, ao computar o produto interno em um espaço de características de dimensão superior (algumas vezes até infinito, conhecido como espaço de Hilbert) (SHAWE-TAYLOR *et al.*, 2004). Isto pode ser exemplificado através da Figura 18, que exhibe dados originais não linearmente separáveis que estão em um espaço de ordem \mathbb{R}^2 . É possível verificar através do gráfico da direita, que os dados são mapeados através de uma função ϕ para o espaço de características de ordem \mathbb{R}^3 e que o problema dual formado é linearmente separável (SVENSÉN; BISHOP, 2007).

Figura 18 – Exemplo de um mapeamento para espaço de características.



Fonte: Schölkopf *et al.* (2002). Modificado pelo autor.

A problemática desta abordagem está associada em computar o mapeamento para este espaço de características. Um exemplo disto, é quando se tem um conjunto de dados $\mathbf{X} \in \mathbb{R}^{m \times n}$, e deseja-se mapear para um espaço de características de ordem maior que n ou até mesmo infinita. Deste modo, o produto interno torna-se computacionalmente inviável e é necessária uma abordagem para contornar este problema. Logo, faz-se o uso de uma técnica que realiza o produto interno no espaço de característica, sem explicitamente conhecer o mapeamento. Este tipo de abordagem é conhecida como o truque do kernel e é realizada através de uma função de kernel que reproduz tal produto interno no espaço de Hilbert. Esta pode ser definida através da seguinte equação:

$$\kappa(\mathbf{x}, \mathbf{x}') = \langle \phi(\mathbf{x}), \phi(\mathbf{x}') \rangle, \quad (4.2)$$

que para todos os vetores $\mathbf{x}, \mathbf{x}' \in \mathbf{X}$ a satisfaça (SCHÖLKOPF *et al.*, 2002; SHAWE-TAYLOR *et al.*, 2004). Além disto, destaca-se que ϕ realiza um mapeamento de \mathbf{X} para o espaço de

características H conforme a equação:

$$\phi : \mathbf{x} \mapsto \phi(\mathbf{x}) \in H \subseteq \mathbb{R}^N. \quad (4.3)$$

Portanto, o desafio ao utilizar este conceito é projetar funções de kernel que sempre consigam realizar o produto interno no espaço de características H de modo que a Equação (4.2) seja verdadeira (SCHÖLKOPF *et al.*, 2002). Há maneiras de verificar se uma função pode ser considerada como kernel, ao constatar a condição necessária e suficiente de que a matriz gramiana \mathbf{K} , seja semidefinida positiva para todas as escolhas de vetores em \mathbf{X} (SVENSÉN; BISHOP, 2007). Sendo assim, dado um conjunto de vetores $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_\ell\}$ a matriz gramiana que possui ordem $\ell \times \ell$ tem seus elementos computados através da equação:

$$\mathbf{K}_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle, \text{ para } i, j = \{1, \dots, \ell\}. \quad (4.4)$$

Destaca-se que pelo fato de se utilizar da mesma Equação (4.2), esta matriz é comumente chamada de matriz de kernel. Como a função de kernel é simétrica em seus argumentos, ou seja, $\kappa(\mathbf{x}, \mathbf{x}') = \kappa(\mathbf{x}', \mathbf{x})$ pode-se exemplificar como a construção da matriz de kernel é realizada através da seguinte equação (SHAWE-TAYLOR *et al.*, 2004):

$$\mathbf{K} = \begin{bmatrix} \langle \phi(\mathbf{x}_1), \phi(\mathbf{x}_1) \rangle & \langle \phi(\mathbf{x}_1), \phi(\mathbf{x}_2) \rangle & \cdots & \langle \phi(\mathbf{x}_1), \phi(\mathbf{x}_\ell) \rangle \\ \langle \phi(\mathbf{x}_2), \phi(\mathbf{x}_1) \rangle & \langle \phi(\mathbf{x}_2), \phi(\mathbf{x}_2) \rangle & \cdots & \langle \phi(\mathbf{x}_2), \phi(\mathbf{x}_\ell) \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \phi(\mathbf{x}_\ell), \phi(\mathbf{x}_1) \rangle & \langle \phi(\mathbf{x}_\ell), \phi(\mathbf{x}_2) \rangle & \cdots & \langle \phi(\mathbf{x}_\ell), \phi(\mathbf{x}_\ell) \rangle \end{bmatrix}. \quad (4.5)$$

Através desta representação, é possível verificar que a matriz de kernel também é simétrica, já que $\mathbf{K}_{ij} = \mathbf{K}_{ji}$. Portanto, uma função de kernel válida produz uma matriz de kernel \mathbf{K} se esta for simétrica e semidefinida positiva (SHAWE-TAYLOR *et al.*, 2004).

A fim de exemplificar o que foi descrito sobre a construção das matrizes de kernel, escolheu-se duas funções de kernel amplamente utilizadas na área do reconhecimento de padrões (SVENSÉN; BISHOP, 2007). A primeira pode ser verificada através de um exemplo específico, para que em sequência possa-se formalizar sua equação geral. Então, ao considerar um espaço de entradas $\mathbf{x} \in \mathbb{R}^2$ pode-se verificar que um mapeamento ϕ para o espaço de características é realizado como se segue (SCHÖLKOPF *et al.*, 2002)

$$\phi : \mathbf{x} = (x_1, x_2) \mapsto \phi(\mathbf{x}) = (x_1^2, x_2^2, \sqrt{2}x_1x_2) \quad (4.6)$$

É possível identificar que a função ϕ recebe dois parâmetros como argumento e retorna três valores. Desta maneira, o vetor de entrada é mapeado para o espaço de característica de ordem superior. Nesta equação ainda é possível verificar que há uma multiplicação de $\sqrt{2}$ em um dos termos de $\phi(\mathbf{x})$. Isto ocorre porque a função ϕ computa de modo não ordenado com substituição, sendo assim necessário a compensação através do escalonamento ao multiplicar por $\sqrt{2}$ (SCHÖLKOPF *et al.*, 2002). Dito isto, pode-se verificar que a abordagem descrita pela Equação (4.6) trata-se de um kernel válido conforme a seguinte expressão:

$$\begin{aligned}
\langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle &= \langle (x_1^2, x_2^2, \sqrt{2}x_1x_2), (y_1^2, y_2^2, \sqrt{2}y_1y_2) \rangle \\
&= x_1^2 \cdot y_1^2 + x_2^2 \cdot y_2^2 + 2 \cdot x_1 \cdot x_2 \cdot y_1 \cdot y_2 \\
&= (x_1 \cdot y_1)^2 + 2 \cdot x_1 \cdot x_2 \cdot y_1 \cdot y_2 + (x_2 \cdot y_2)^2, \\
&= (x_1 \cdot y_1 + x_2 \cdot y_2)^2 \\
&= \langle \mathbf{x}, \mathbf{y} \rangle^2
\end{aligned} \tag{4.7}$$

e portanto o produto interno das projeções no espaço de características pode ser computado sem ter que conhecer de fato suas coordenadas (SHAWE-TAYLOR *et al.*, 2004). Esta abordagem pode ser generalizada para qualquer ordem d de polinômios, e por isto a seguinte equação:

$$\kappa(\mathbf{x}, \mathbf{x}') = (\langle \mathbf{x}, \mathbf{x}' \rangle + c)^d, \tag{4.8}$$

é comumente conhecida como função de kernel polinomial. Nesta é possível verificar a presença do parâmetro livre $c > 0$ para garantir que seu intervalo numérico esteja dentro de algum limite. Comumente este parâmetro é dependente da faixa de valores do conjunto \mathbf{X} (SHAWE-TAYLOR *et al.*, 2004).

Uma abordagem comum em diversas pesquisas é a construção de funções de kernel complexas a partir de uma ou mais funções de kernel simples. Isto é possível, pois a matriz computada por essas funções pode ser validada através da propriedade semidefinida finitamente positiva mencionada anteriormente. Ou seja, operações realizadas nas funções e, consequentemente nas matrizes, produzirão matrizes de kernel enquanto forem matrizes simétricas e semidefinidas positivas, portanto, são capazes de mapear os dados de entrada para o espaço de características de Hilbert (SHAWE-TAYLOR *et al.*, 2004). Nesse contexto, utiliza-se na presente dissertação, a função de kernel gaussiana representada pela seguinte equação:

$$\kappa(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right), \tag{4.9}$$

que se enquadra em uma classe chamada de funções de base radial (ou *radial basis function* (RBF)), que são dependentes da magnitude da distância entre seus argumentos. Afirma-se ainda que a distância utilizada em seu cálculo não é limitada apenas a euclidiana (SVENSÉN; BISHOP, 2007).

Ao utilizar a função de kernel RBF é possível constatar algumas informações importantes. A primeira que a imagem dos pontos no espaço de características tem norma unitária quando $\kappa(\mathbf{x}, \mathbf{x})$. Como todos os produtos internos dos pontos mapeados são positivos, é possível fazer com que todas as imagens estejam em um único ortante (SCHÖLKOPF *et al.*, 2002). Outra propriedade importante é que o kernel RBF é um kernel Polinomial de ordem infinita. Ou seja, sua aplicação reproduz um espaço de características de dimensão infinita (SHAWE-TAYLOR *et al.*, 2004). Isso é verificado, pois, a função exponencial pode ser aproximada através da expansão da série de Taylor conforme:

$$\exp(x) = \sum_{i=0}^{\infty} \frac{1}{i!} x^i. \quad (4.10)$$

Pode-se realizar a verificação de que uma função kernel RBF é um kernel válido, através das seguintes propriedades: $\kappa(\mathbf{x}, \mathbf{x}') = a\kappa_1(\mathbf{x}, \mathbf{x}')$; $\kappa(\mathbf{x}, \mathbf{x}') = \exp(\kappa_1(\mathbf{x}, \mathbf{x}'))$ (SHAWE-TAYLOR *et al.*, 2004). Considerando κ_1 uma função de kernel válida e $a \in \mathbb{R}^+$. Sendo assim, a validação da função RBF é exibida através da seguinte expressão:

$$\begin{aligned} \kappa(\mathbf{x}, \mathbf{x}') &= \exp(\langle \mathbf{x}, \mathbf{x}' \rangle / \sigma^2) \\ \frac{\kappa(\mathbf{x}, \mathbf{x}')}{\sqrt{\kappa(\mathbf{x}, \mathbf{x}) \cdot \kappa(\mathbf{x}', \mathbf{x}')}} &= \frac{\exp(\langle \mathbf{x}, \mathbf{x}' \rangle / \sigma^2)}{\sqrt{\exp(\langle \mathbf{x}, \mathbf{x} \rangle / \sigma^2) \cdot \exp(\langle \mathbf{x}', \mathbf{x}' \rangle / \sigma^2)}} \\ &= \frac{\exp(\langle \mathbf{x}, \mathbf{x}' \rangle / \sigma^2)}{\exp(\langle \mathbf{x}, \mathbf{x} \rangle / \sigma^2 + \langle \mathbf{x}', \mathbf{x}' \rangle / \sigma^2)^{\frac{1}{2}}} \\ &= \exp\left(\frac{\langle \mathbf{x}, \mathbf{x}' \rangle}{\sigma^2}\right) \cdot \exp\left(-\frac{\langle \mathbf{x}, \mathbf{x} \rangle}{2\sigma^2} - \frac{\langle \mathbf{x}', \mathbf{x}' \rangle}{2\sigma^2}\right) \\ &= \exp\left(\frac{\langle \mathbf{x}, \mathbf{x}' \rangle}{\sigma^2} - \frac{\langle \mathbf{x}, \mathbf{x} \rangle}{2\sigma^2} - \frac{\langle \mathbf{x}', \mathbf{x}' \rangle}{2\sigma^2}\right) \\ &= \exp\left(\frac{2\langle \mathbf{x}, \mathbf{x}' \rangle - \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}', \mathbf{x}' \rangle}{2\sigma^2}\right) \\ &= \exp\left(-\frac{\langle \mathbf{x} - \mathbf{x}', \mathbf{x} - \mathbf{x}' \rangle}{2\sigma^2}\right) \\ &= \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right), \end{aligned} \quad (4.11)$$

que faz inicialmente uma normalização, transformando-o em um kernel normalizado.

4.2 Esparsificação de Modelos

Com a explicação da utilização de matrizes kernel, faz-se agora uma descrição da problemática associada ao utilizar estas matrizes para parametrização. Como dito anteriormente, os sinais de EEG são analisados em um curto segmento do sinal chamado de época. Considerando uma época com $2s$, um dispositivo possuindo uma quantidade K de canais e que faz aquisição dos sinais com uma taxa de amostragem (SR) em Hz , tem-se para cada época, um segmento de sinal com $N = 2 \cdot SR$ amostras para apenas um canal. Considerando os K canais a cada época tem-se uma matriz de dados $\mathbf{X}_{\text{época}} \in \mathbb{R}^{N \times K}$.

Ao enviar esta matriz $\mathbf{X}_{\text{época}}$ para a função que estima a matriz de kernel, tem-se como resposta uma matriz $\mathbf{K}_{\text{época}} \in \mathbb{R}^{N \times N}$. Destaca-se que por conta desta matriz ter uma dimensão muito alta, é exigido um alto desempenho computacional para estimá-la. Além disto, ao vetorizar a matriz, tem-se um vetor de características com ordem $\mathbb{R}^{N \cdot (N+1)/2 \times 1}$. Portanto, utilizar um vetor com esta quantidade de elementos, faz com que o tempo de estimação dos modelos na etapa de classificação, seja altíssimo, e dependendo do método, é computacionalmente inviável.

Pode-se contornar o problema exposto ao trabalhar-se com matrizes de kernel compactas, utilizando métodos de imposição da esparsificação da matriz de dados. Este trata-se de uma seleção de vetores na matriz, de modo que seja criado um dicionário de vetores relevantes (HAN *et al.*, 2011). Como não há possibilidade de aplicar esta seleção em uma matriz de kernel já estimada, faz-se a realização da seleção na massa de dados $\mathbf{X}_{\text{época}}$. O fato disto, é que a alteração dos elementos da matriz de kernel resulta na perda das propriedades de simetria e faz com que a mesma não seja mais positiva definida (SUYKENS; VANDEWALLE, 1999).

Fundamentalmente, a escolha dos vetores pode ser realizada de modo aleatório, porém, para compor um dicionário de vetores relevantes, o presente trabalho descreve sobre dois métodos baseados na imposição de esparsificação do modelo. O primeiro utiliza conceito de quantização vetorial, e o segundo através de critério da maximização de entropia computada a partir da matriz de kernel compacta.

4.2.1 Esparsificação através de quantização vetorial

No presente trabalho, escolheu-se o método de quantização vetorial através de um método de clusterização. Estes métodos são extensivamente utilizados em AM para resolver problemas associados ao reconhecimento de padrões. A tarefa associada ao método de agrupa-

mento é não supervisionada, ou seja, não há informações sobre as classes no conjunto de dados, portanto, realiza-se uma análise nos dados de modo a descobrir estruturas e similaridades para que agrupamentos sejam formados (SVENSÉN; BISHOP, 2007; WEBB, 2003).

Há diversos métodos de agrupamentos e são fundamentalmente baseados na exploração em determinado conjunto de dados, de modo a se encontrar grupos naturais. Caso sejam encontrados grupos compactos, pode-se reduzir a informação do conjunto original para uma quantidade menor de grupos, em alguns casos um grupo pode ser representado por uma única amostra. Para o presente trabalho escolheu-se o método k -Médias, que faz um particionamento do conjunto de dados em k grupos não sobrepostos, de modo a minimizar o somatório das distâncias quadráticas nas partições (WEBB, 2003).

Considerando o conjunto de dados mencionado anteriormente $\mathbf{X} \in \mathbb{R}^{N \times K}$, a aplicação do método do k -Médias é realizada como exibido através do Algoritmo 1. Inicialmente, faz-se uma seleção aleatória de k protótipos iniciais representados por $\mathbf{W} \in \mathbb{R}^{k \times K}$ a partir do conjunto de dados \mathbf{X} . Em sequência, a i -ésima partição é definida como

$$\mathbb{V}_i = \{\mathbf{x} \in \mathbb{R}^K \mid \|\mathbf{x} - \mathbf{w}_i\| < \|\mathbf{x} - \mathbf{w}_j\|, \forall i \neq j\}, \quad (4.12)$$

que é possível identificar que os dados em \mathbf{X} são associados aos protótipos que possuem menor distância euclidiana (WEBB, 2003). Destaca-se ainda que $i = j = \{1, \dots, k\}$.

Algoritmo 1: k -Médias.

Input: $\mathbf{X}_{\text{época}} \in \mathbb{R}^{(N \times K)}$, k .

Output: Conjunto particionado $\mathbf{E} \in \mathbb{R}^{k \times K}$.

Início;

Posicionamento de k protótipos;

while Não atingir a convergência;

do

 Determinar a partição \mathbb{V}_i do protótipo \mathbf{w}_i pela Equação (4.12);

 Novo posicionamento de protótipos $\mathbf{w}_i = \frac{1}{N} \sum_{\mathbf{x} \in \mathbb{V}_i} \mathbf{x}$;

 Cálculo de SSD pela Equação 4.13;

end

Após o particionamento dos dados, computa-se a soma das distâncias quadráticas entre cada elemento na partição e protótipo associado através da seguinte equação:

$$SSD = \sum_{i=1}^k \sum_{\forall \mathbf{x} \in \mathbb{V}_i} \|\mathbf{x} - \mathbf{w}_i\|^2. \quad (4.13)$$

Em sequência, faz-se a atualização dos protótipos, modificados pelo valor de ponto médio de cada partição. Como este método é exaustivo, o procedimento é repetido até que com a partição

dos dados não produza uma diminuição da soma das distâncias quadráticas (WEBB, 2003). Por fim, para compor o dicionário de vetores relevantes a partir do método k -Médias, escolheu-se para este trabalho duas abordagens: seleção dos k protótipos finais; seleção dos k elementos mais próximos aos protótipos de cada partição. Desta maneira, ao aplicar o método no conjunto de dados $\mathbf{X} \in \mathbb{R}^{N \times K}$ tem-se um dicionário de vetores relevantes $\mathbf{E} \in \mathbb{R}^{k \times K}$.

4.2.2 Esparsificação pelo critério de entropia de Renyi

O segundo método para a composição do dicionário de vetores relevantes, é realizado através da imposição de esparsificação pelo critério de maximização da entropia quadrática de Renyi. Diferente do método anterior, este se utiliza a matriz de kernel para o cálculo da entropia, e através de um processo exaustivo, busca a seleção de m vetores do conjunto de dados que proporcionam um aumento no valor da entropia quadrática de Renyi (SANTOS; BARRETO, 2018).

Considerando o conjunto de dados original de ordem $\mathbf{X} \in \mathbb{R}^{N \times K}$, o procedimento para aplicação do método proposto, pode ser visualizado através do Algoritmo 2. Inicialmente deve-se definir o valor de m que representa a quantidade de vetores a serem selecionados. Em seguida, o conjunto \mathbf{X} é dividido em dois subconjuntos. O primeiro $\mathbf{E} \in \mathbb{R}^{m \times K}$, possui m vetores selecionados aleatoriamente, e o segundo $\mathbf{Z} \in \mathbb{R}^{N-m \times K}$ é composto pelos vetores restantes. Com estes definidos, faz-se iterativamente a troca de um vetor linha presente em \mathbf{E} por um vetor linha em \mathbf{Z} . Para cada iteração computa-se a entropia quadrática de Renyi e caso seu valor tenha uma magnitude superior ao maior valor de entropia já calculado anteriormente, mantém-se a troca entre vetores. Esse processo é repetido até que o número máximo de iterações seja atingido (SANTOS; BARRETO, 2018).

Vale ressaltar que a entropia quadrática de Renyi é dada pela equação:

$$H = -\log \int p(\mathbf{x})^2 d\mathbf{x}, \quad (4.14)$$

a qual pode ser aproximada para a seguinte equação:

$$H = -\log \left(\frac{1}{m^2} \cdot \mathbf{1}_m^T \cdot \Sigma \cdot \mathbf{1}_m \right). \quad (4.15)$$

Nesta Equação, destaca-se que Σ representa a matriz de kernel à ser avaliada e $\mathbf{1}_m$ é um vetor unitário de ordem m (SANTOS; BARRETO, 2018). Similarmente ao método anterior, ao aplicar este no conjunto de dados \mathbf{X} , produz o dicionário de vetores relevantes $\mathbf{E} \in \mathbb{R}^{m \times K}$.

Algoritmo 2: Esparsificação pela entropia.

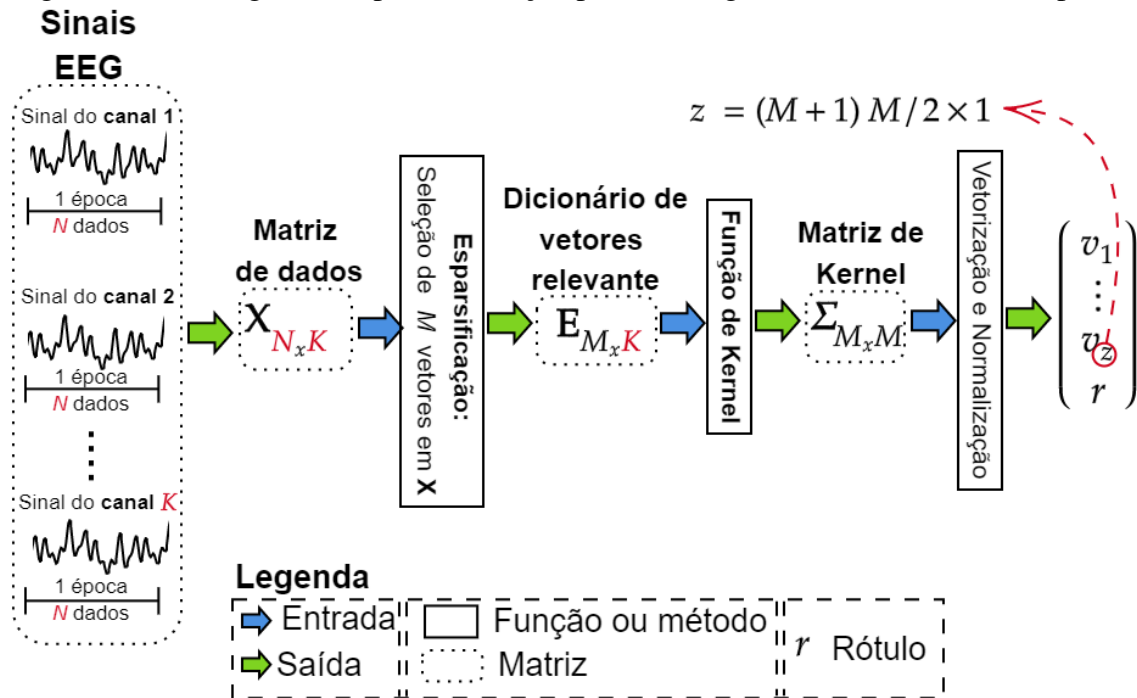
Input: $\mathbf{X} \in \mathbb{R}^{N \times K}$, m , Quantidade de iterações.
Output: Subconjunto \mathbf{E}
 $\mathbf{E} \leftarrow \{\mathbf{x}_m \text{ aleatórios}\};$
 $\mathbf{Z} \leftarrow \{\mathbf{x}_{(N-m) \text{ restantes}}\};$
 $\mathbf{K} \leftarrow$ Computa-se matriz de kernel em \mathbf{E} ;
 $E1 \leftarrow$ Calcula-se entropia de Renyi em \mathbf{K} ;
while Não atingir número máximo de iterações;
do
 Troca-se aleatoriamente um padrão de \mathbf{E} por \mathbf{Z} ;
 Computa-se matriz de kernel em \mathbf{E} ;
 $E2 \leftarrow$ Calcula-se entropia de Renyi em \mathbf{K} ;
 if $E2 > E1$ **then**
 $E1 \leftarrow E2$;
 Troca é efetivada;
 end
end

Após a explicação da composição de matrizes de kernel e a aplicação do método de esparsificação, faz-se necessário a descrição de como estes são utilizados em conjunto de modo a parametrizar o sinal de EEG. Faz-se necessário lembrar que no presente trabalho, duas análises são utilizadas: intercanais e intracanal. Portanto, considerando apenas uma época a ser analisada pela abordagem intercanais, o procedimento genérico de parametrização pela matriz de kernel em conjunto com a esparsificação, pode ser visualizada através da Figura 19.

Nesta Figura é possível verificar que inicialmente os sinais presentes em todos os canais são analisados ao compor a matriz $\mathbf{X} \in \mathbb{R}^{N \times K}$. Em sequência, aplica-se o método de esparsificação selecionado de modo a construir o dicionário com os M vetores relevantes. Vale ressaltar que $M = k$ quando se utiliza o método k -Médias e $M = m$ quando se aplica o método de maximização da entropia de Renyi. Após a composição do dicionário de vetores relevantes $\mathbf{E} \in \mathbb{R}^{M \times K}$, estima-se a matriz de kernel $\Sigma \in \mathbb{R}^{M \times M}$, através da função de kernel escolhida e faz-se o procedimento de normalização e vetorização idêntico ao utilizado no método de covariância e por fim, o vetor de atributos gerado é rotulado. Destaca-se que o procedimento exposto pelo fluxograma, é repetido para todas as possíveis épocas, concatenando os vetores de atributos rotulados pelas colunas.

A abordagem intracanal da utilização dos métodos propostos pelo presente trabalho é explicada pelo fluxograma presente na Figura 20. Este representa a parametrização pela abordagem intracanal para apenas uma época. Pode-se perceber a diferença da abordagem

Figura 19 – Fluxograma de parametrização pela abordagem intercanais em uma época.



Fonte: Elaborado pelo Autor.

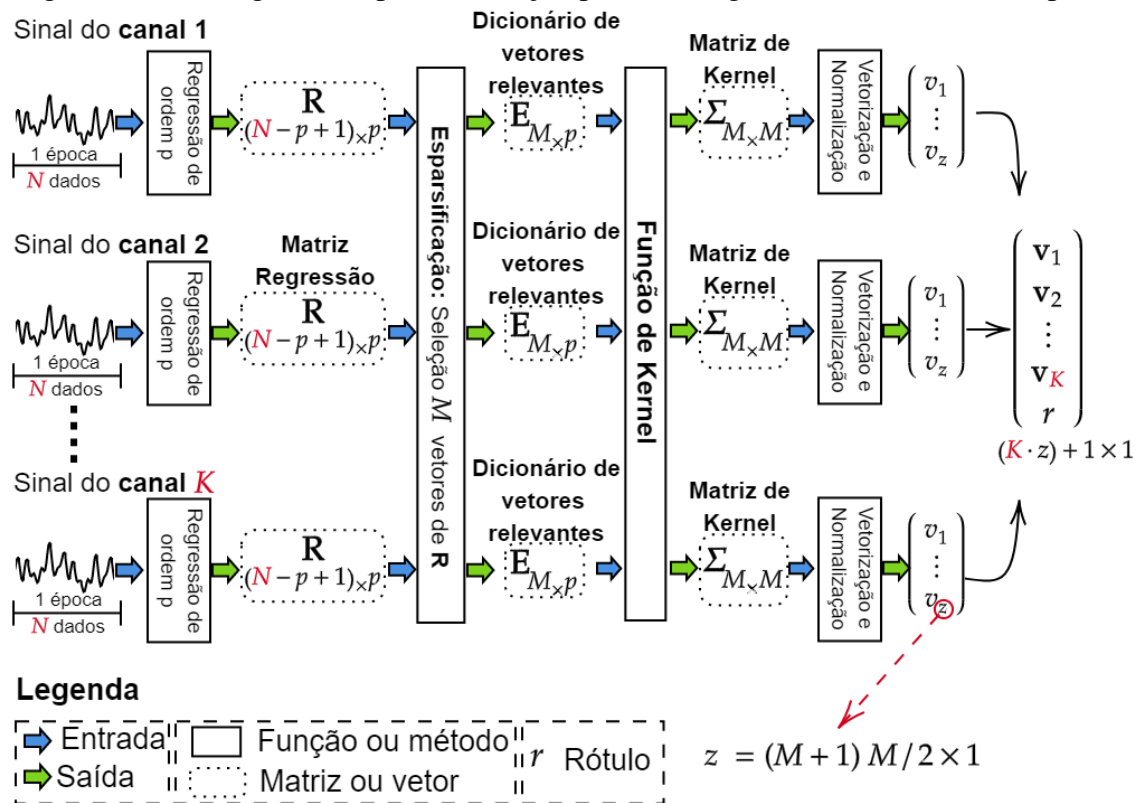
intercanais, pois, esta faz análises do sinal em cada canal individualmente. Ou seja, para cada canal tem-se um vetor $\mathbf{x}_i \in \mathbb{R}^{N \times K}$ que representa as N amostras naquele canal. Este vetor é utilizado para compor a matriz de regressores $\mathbf{R}_i \in \mathbb{R}^{N-p+1 \times p}$ cuja estimação é realizada pela Equação (3.7) já descrita anteriormente. Em seguida, aplica-se o método de esparificação escolhido em \mathbf{R}_i , de modo a selecionar os M ($M = k$ para k -Médias e $M = m$ para entropia de Renyi) vetores que compõe o dicionário de vetores relevantes $\mathbf{E}_i \in \mathbb{R}^{M \times p}$. Com isto, pode-se então aplicar a função de kernel escolhida para construir a matriz $\Sigma_i \in \mathbb{R}^{M \times M}$, que pode ser em seguida vetorizada e normalizada \mathbf{v}_i .

O procedimento então é repetido para todos os K canais para que todos os vetores $\mathbf{v}_i | i = \{1, 2, 3 \dots K\}$ sejam concatenados pela linha, formando assim o vetor final de atributos rotulado de ordem $(K \cdot ([M + 1] \cdot M/2)) + 1 \times 1$. Igualmente para a abordagem intercanais, o procedimento descrito é repetido para todas as épocas disponíveis, realizando a concatenação dos vetores de atributos rotulados pelas colunas.

4.3 Classificadores

A presente seção aborda sobre os algoritmos de AM utilizados no trabalho para a etapa de classificação dos vetores de atributos rotulados. Para o presente trabalho, foram

Figura 20 – Fluxograma de parametrização pela abordagem intracanal em uma época.



Fonte: Elaborado pelo Autor.

escolhidos dois métodos de classificação lineares. O argumento para esta decisão reside no fato de que se classificadores lineares apresentassem bom desempenho ao utilizar os métodos de extração não lineares propostos na presente dissertação, muito provavelmente classificadores não lineares também teriam bom desempenho. Além disso, os classificadores escolhidos não possuem hiperparâmetros, o que facilita sobremaneira o processo de treinamento. Tal escolha também é justificada pelo fato de os métodos de parametrização avaliados na dissertação já possuírem diversos hiperparâmetros a serem sintonizados. Não seria conveniente introduzir ainda mais hiperparâmetros durante o processo de classificação.

4.3.1 Método dos mínimos quadrados

O método dos mínimos quadrados ordinários, *ordinary least squares* (OLS), é um método baseado em aprendizado supervisionado que é amplamente utilizado como regressão. Ou seja, para prever um valor contínuo a partir de um vetor de entradas. Essa relação pode ser expressada através da seguinte equação:

$$\mathbf{y}_i(\mathbf{x}) = \mathbf{w}_i^T \mathbf{x} = \langle \mathbf{w}_i, \mathbf{x} \rangle, \quad (4.16)$$

na qual é possível verificar que é calculado a similaridade entre \mathbf{x} e \mathbf{w} (SHAWE-TAYLOR *et al.*, 2004). Portanto, este método pode ser utilizado para resolver problemas de classificação através desta similaridade. Esta Equação representa apenas um dos possíveis resultados no universo de classes presentes no problema em questão então, pode-se reescrevê-la como $\mathbf{Y} = \mathbf{W}\mathbf{X}$, que agrupa os vetores de saída em uma matriz. Portanto, para um problema de classificação com C diferentes classes, um conjunto $\mathbf{X} \in \mathbb{R}^{P \times A}$ com um modelo $\mathbf{W} \in \mathbb{R}^{P \times C}$ já estimado, a saída produzida pela equação matricial é $\mathbf{Y} \in \mathbb{R}^{C \times A}$ (SVENSÉN; BISHOP, 2007).

Na etapa de treinamento deste métodos, as matrizes \mathbf{X} e \mathbf{Y} são conhecidas, pois, um conjunto de treinamento é composto por $\{(\mathbf{x}_1, \mathbf{y}_1) \cdots (\mathbf{x}_P, \mathbf{y}_P)\}$. É necessário, portanto a estimação do modelo \mathbf{W} através do critério de minimização da soma dos erros quadráticos representado pela Equação 4.17 (SVENSÉN; BISHOP, 2007).

$$\begin{aligned} \|\varepsilon\|_2^2 &= \|\mathbf{Y} - \mathbf{X}\mathbf{W}\|_2^2 \\ &= (\mathbf{X}\mathbf{W} - \mathbf{Y})^T (\mathbf{X}\mathbf{W} - \mathbf{Y}) \\ &= \mathbf{W}^T \mathbf{X}^T \mathbf{X} \mathbf{W} - 2\mathbf{W}^T \mathbf{X}^T \mathbf{Y} + \mathbf{Y}^T \mathbf{Y}. \end{aligned} \quad (4.17)$$

O procedimento de obtenção do estimador dos parâmetro do modelo é desenvolvido como

$$\begin{aligned} \mathbf{0} &= \frac{\partial \|\varepsilon\|_2^2}{\partial \mathbf{W}} \\ \mathbf{0} &= \frac{\partial (\mathbf{W}^T \mathbf{X}^T \mathbf{X} \mathbf{W} - 2\mathbf{W}^T \mathbf{X}^T \mathbf{Y} + \mathbf{Y}^T \mathbf{Y})}{\partial \mathbf{W}} \\ \mathbf{0} &= 2\mathbf{X}^T \mathbf{X} \mathbf{W} - 2\mathbf{X}^T \mathbf{Y} + \mathbf{0} \end{aligned} \quad (4.18)$$

$$2\mathbf{X}^T \mathbf{X} \mathbf{W} = 2\mathbf{X}^T \mathbf{Y}$$

$$\mathbf{W} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y},$$

em que é possível identificar que se aplica a derivada na função com respeito a \mathbf{W} e iguala-se a zero. Com isto, se a inversa de $\mathbf{X}^T \mathbf{X}$ existir, a solução do problema dos mínimos quadrados é realizada através da última linha na Equação 4.18 (SVENSÉN; BISHOP, 2007).

Por fim, na etapa de teste, seja uma amostra $\mathbf{x}_{teste} \in \mathbb{R}^{P \times 1}$ a ser avaliada a ser avaliado, faz-se o cálculo $\mathbf{y}_{teste} = \mathbf{W}^T \mathbf{x}_{teste}$. Com isto, forma-se o vetor $\mathbf{y}_{teste} \in \mathbb{R}^{C \times 1}$ que possui os C elementos com as medidas de similaridades. O critério de discriminação para a amostra \mathbf{x}_{teste} é realizado conforme a seguinte equação:

$$j^* = \arg \max \{y_j\} \forall j, \quad (4.19)$$

que busca o índice do vetor com o maior valor y_{teste} e faz-se a rotulação da classe associada a esse índice (SVENSÉN; BISHOP, 2007).

4.3.2 Classificador Linear Gaussiano

Descreve-se nesta seção o CLG que é baseado na densidade de probabilidade gaussiana multivariada e possui um critério de decisão através de dissimilaridade, ao computar a distância de Mahalanobis. Ao considerar um problema de classificação com uma quantidade C de classes e um conjunto A de pares $\{\mathbf{x}_a, y_a\}_{a=1}^A$, é de interesse saber à qual classe uma nova amostra pertence. Neste caso, $\mathbf{x}_a \in \mathbb{R}^P$ expressa o a -ésimo vetor de atributo, y_a seu respectivo rótulo e $A = \sum_{i=1}^C a_i$ a quantidade de amostras.

Pode-se supor que todos os padrões de cada classe são equiprováveis, portanto, a probabilidade *a priori* que uma observação aleatória do conjunto de dados pertence a uma das classes é $p(y_i) = a_i/A$. Para desempenhar a tarefa de classificação proposta, utiliza-se da probabilidade *a posteriori* $p(y_i|\mathbf{x}_a)$, ou seja, a probabilidade de y_i ocorrer dado \mathbf{x}_a . Isto pode ser desempenhado através do teorema de Bayes expressado pela seguinte equação:

$$p(y_i|\mathbf{x}_a) = \frac{p(y_i)p(\mathbf{x}_a|y_i)}{p(\mathbf{x}_a)}, \quad (4.20)$$

que faz uma relação entre a probabilidade *a priori* e a função de densidade $p(\mathbf{x}_a|y_i)$. Esta por sua vez é comumente conhecida como função de verossimilhança e é um modelo probabilístico que expressa como os dados estão organizados (JAMES, 2021; WEBB, 2003).

Um modelo utilizado comumente para a função de verossimilhança é a densidade normal multivariada exibida através da equação:

$$p(\mathbf{x}_a|y_i) = \frac{1}{(2\pi)^{P/2}|\mathbf{C}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}_a - \mathbf{m}_i)^T \mathbf{C}^{-1}(\mathbf{x}_a - \mathbf{m}_i)\right), \quad (4.21)$$

em que \mathbf{m}_i é o vetor médio do subconjunto \mathbf{x} que pertence à classe y_i e \mathbf{C} é a matriz de covariância de \mathbf{X} (JAMES, 2021).

A partir da Equação (4.20), pode-se derivar diversos critérios de tomada de decisão de tal maneira que faz-se uma atribuição de um padrão à uma classe de acordo com a seguinte equação:

$$\text{Atribuir } \mathbf{x}_a \text{ à classe } y_j, \text{ se } g_j(\mathbf{x}_a) > g_i(\mathbf{x}_a), \forall i \neq j. \quad (4.22)$$

Com isto, pode-se notar que se $g_i(\mathbf{x}_a) = p(y_i|\mathbf{x}_a)$ o termo $p(\mathbf{x}_a)$ não influencia numa possível tomada de decisão e pode ser cancelado.

Destaca-se ainda que a função $g_i(\mathbf{x}_a)$ é qualquer função que represente um valor escalar que define uma pertinência do padrão \mathbf{x}_a à classe y_i . Uma abordagem extensivamente utilizada é exibida através da Equação 4.23.

$$\begin{aligned}
 g_i(\mathbf{x}_a) &= \ln[p(w_i|\mathbf{x}_a)] \\
 &= \ln[p(w_i)] + \ln[p(\mathbf{x}_a|w_i)] \\
 &= \ln[p(w_i)] + \ln \left[\frac{1}{(2\pi)^{P/2}|\mathbf{C}|^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{x}_a - \mathbf{m}_i)^T \mathbf{C}^{-1}(\mathbf{x}_a - \mathbf{m}_i) \right\} \right] \\
 &= -\frac{1}{2}(\mathbf{x}_a - \mathbf{m}_i)^T \mathbf{C}^{-1}(\mathbf{x}_a - \mathbf{m}_i) - \frac{1}{2} \ln[|\mathbf{C}|] + \ln[p(w_i)]
 \end{aligned} \tag{4.23}$$

Como dito anteriormente, uma suposição comum é de que as classes são equiprováveis e portanto o termo $\ln[p(w_i)]$ pode ser desconsiderado. Além disto, como a matriz de covariância é igual para todas as classes, pode-se também eliminar o termo $-\frac{1}{2} \ln[|\mathbf{C}|]$, produzindo assim a equação:

$$g_i(\mathbf{x}_a) = Q_i(\mathbf{x}_a) = (\mathbf{x}_a - \mathbf{m}_i)^T \mathbf{C}^{-1}(\mathbf{x}_a - \mathbf{m}_i). \tag{4.24}$$

Esta função discriminante é conhecida como distância de Mahalanobis quadrática e para este caso, pode-se utilizar como critério de decisão o valor mínimo das distâncias calculadas entre o vetor \mathbf{x}_a e todos os vetores médios, como pode ser visto através da seguinte equação:

$$\text{Atribuir } \mathbf{x}_a \text{ à classe } y_j, \text{ se } Q_j(\mathbf{x}_a) < Q_i(\mathbf{x}_a), \forall i \neq j. \tag{4.25}$$

Para o presente trabalho, ainda foi utilizado uma abordagem para minimizar a possibilidade de a matriz de covariância ser não invertível. Esta abordagem faz a estimação das matrizes de covariância para cada uma das classes, porém, realizando uma multiplicação com a respectiva probabilidade *a priori*. Desta maneira a matriz de covariância final, chamada agregada, é composta pela média ponderada entre probabilidade *a priori* e matriz de covariância, como pode ser visualizado através da seguinte equação:

$$\mathbf{C}_{\text{agregada}} = \left(\frac{a_1}{A}\right) \mathbf{C}_1 + \left(\frac{a_2}{A}\right) \mathbf{C}_2 + \dots + \left(\frac{a_C}{A}\right) \mathbf{C}_C = \sum_{i=1}^C p(y_i) \cdot \mathbf{C}_i \tag{4.26}$$

4.4 Considerações Finais

No decorrer do presente capítulo foram descritas as formalizações matemáticas para cada um dos métodos de parametrização propostos pela presente dissertação. Além disto, explicam-se por fluxogramas como os métodos são aplicados em curtos segmentos do sinal de

EEG. Por fim, o capítulo é finalizado pela descrição dos classificadores utilizados para estimar os modelos de predição dos problemas de classificação.

5 METODOLOGIA DOS EXPERIMENTOS

No presente capítulo descreve-se a metodologia para aplicação dos métodos de parametrização avaliados nesta dissertação. Faz-se inicialmente uma descrição sobre os conjuntos de dados escolhidos, e em seguida explica-se através de algoritmos como cada um dos métodos foi utilizado.

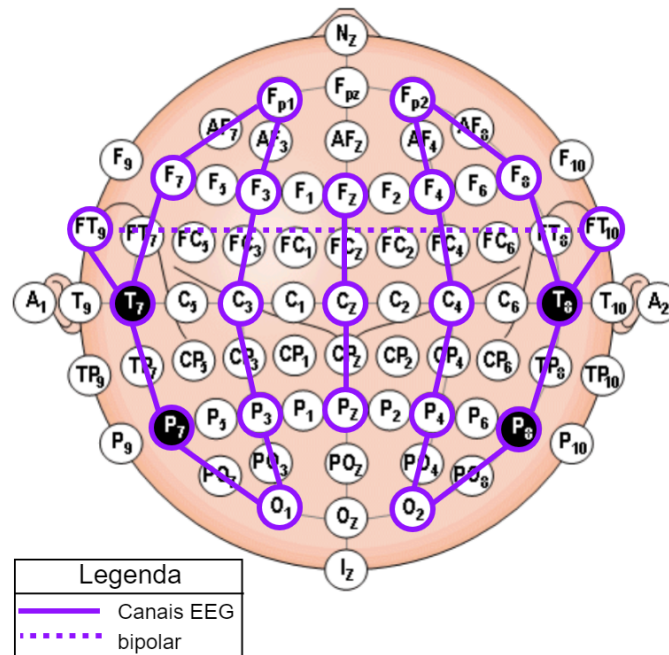
5.1 Descrição do primeiro conjunto de dados

O primeiro conjunto de dados utilizado para testar a parametrização do sinal de EEG pelos métodos propostos na presente dissertação, foi o **CHB-MIT Scalp EEG**. Este é um conjunto de dados de livre acesso e possui dados de pacientes pediátricos com convulsões intratáveis (SHOEB, 2010). Os pacientes analisados pelo autor do conjunto de dados foram monitorados durante diversos dias e sem algum tipo de medicação preventiva contra crises epiléticas, pois são possíveis candidatos à intervenção cirúrgica. O conjunto de dados é composto por 23 casos, que contém os arquivos de 22 pacientes, isto porque os casos **chb01** e **chb21** foram obtidos do mesmo paciente, com uma diferença de um ano e meio. Dos 22 pacientes, 5 eram do sexo masculino e 17 do sexo feminino e suas idades variavam entre 1,5 até 19 anos. O autor relata ainda que após um ano do último experimento, um vigésimo quarto caso foi adicionado, porém este não possui informações sobre sexo e idade (SHOEB, 2009).

A aquisição dos sinais EEG de cada paciente foi realizada com uma taxa de amostragem de 256Hz e os dados possuem resolução de 16 bits. Durante a aquisição, foi realizado o posicionamento de eletrodos do couro cabeludo de acordo com o padrão 10 – 20 ilustrado pela Figura 21. Cada caso possui entre nove a 42 arquivos com o sinal contínuo gravado, porém destes arquivos apenas alguns possuem dados convulsivos. Além disso, a maioria das aquisições foi realizada utilizando 23 canais, porém em algumas situações utilizou-se até 26 canais com outros tipos de sinais como ECG, ou estímulo do nervo vago (do inglês *vagal nerve stimulus* – VNS) (SHOEB, 2009).

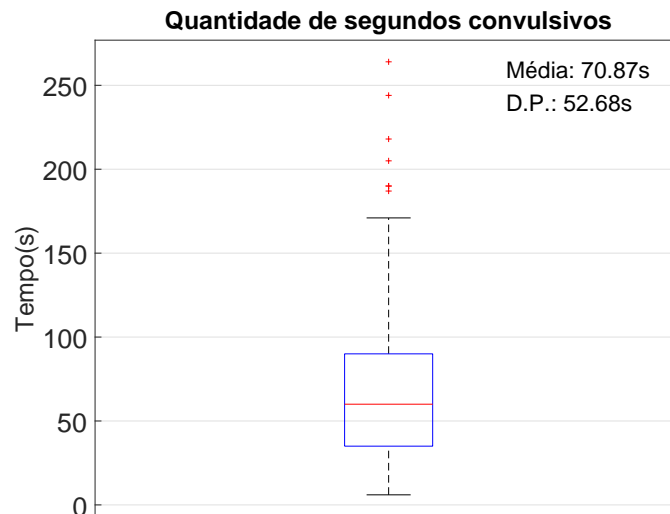
Outra característica importante deste conjunto de dados, é que a maioria dos arquivos possui informações de um monitoramento com duração de uma hora (com exceção daqueles em que a aquisição teve duração de quatro horas). Deste tempo total, apenas poucos segundos são de sinais convulsivos, portanto, verifica-se a presença de um problema que possui classes desbalanceadas. De modo a exibir graficamente a duração dos sinais convulsivos de todos os

Figura 21 – Padrão 10-20 utilizado no conjunto de dados.



Fonte: Acharya *et al.* (2016). Modificado pelo Autor.

Figura 22 – Boxplot com todas as quantidades de períodos convulsivos.

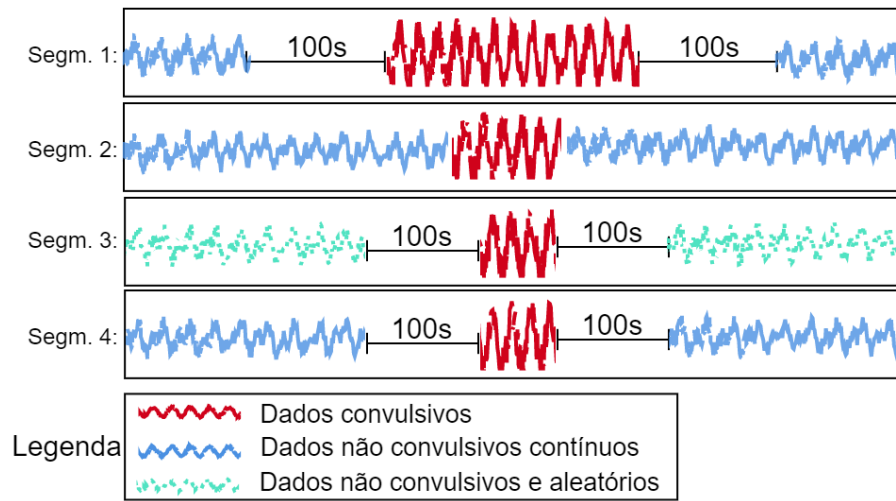


Fonte: Elaborado pelo Autor.

arquivos, utilizou-se o *boxplot* ilustrado pela Figura 22.

Como descrito, o conjunto de dados possui certas despadronizações e de modo a corrigir algumas destas, definiu-se inicialmente que apenas os 23 canais destacados pela Figura 21 são utilizados no presente trabalho. Outra escolha realizada, faz-se a seleção apenas dos arquivos que possuem crises epiléticas identificadas. Por último, foram feitas algumas segmentações do

Figura 23 – Segmentações utilizadas para a etapa de treinamento.



Fonte: Elaborado pelo Autor.

signal completo para a etapa de treinamento dos classificadores. Isto foi realizado selecionando os dados não convulsivos, na tentativa de reduzir um possível efeito transitório dos períodos pré ou pós-convulsivos, bem como na tentativa de balancear a quantidade de classes. A hipótese desta segmentação foi feita segundo o trabalho de OLIVEIRA (2021), que propôs sete diferentes maneiras para tal. Destas, o autor pôde identificar quatro segmentações que aumentaram o desempenho do método dos mínimos quadrados na etapa de classificação, sendo estas escolhidas para o presente trabalho.

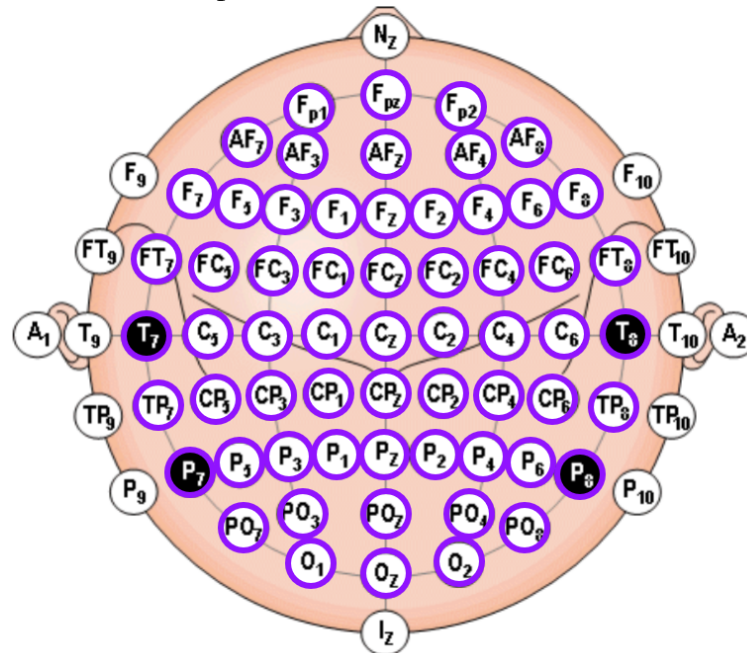
As hipóteses de segmentações escolhidas podem ser ilustradas pela Figura 23, que apresenta um esboço de como é realizada cada segmentação. Os sinais destacados em vermelho referem-se aos períodos convulsivos, os em azul são referentes aos sinais não convulsivos e os representados pela cor verde são sinais não convulsivos, porém, selecionados de maneira aleatória. Além disto, pode-se explicar que a segmentação de número um utiliza de uma mesma quantidade de sinais convulsivos e não convulsivos. Isto se difere das demais, pois, nestas a quantidade de dados não convulsivos é de três vezes a quantidade de dados convulsivos.

5.2 Descrição do segundo conjunto de dados

O segundo conjunto de dados utilizado no presente trabalho é proveniente de um estudo realizado de modo a identificar a predisposição genética ao alcoolismo em pessoas, a partir de sinais de EEG. O conjunto foi composto de aquisições do sinal em uma taxa de amostragem de $256Hz$ através de 64 eletrodos posicionados no couro cabeludo pelo sistema 10 –

20, ilustrado pela Figura 24. Diferente do conjunto de dados anterior, os eletrodos não são em modo bipolar, portanto, cada canal é um respectivo eletrodo com referência ao comum. A análise foi realizada em 122 pessoas que estavam sujeitas a determinado estímulo visual. Este estímulo tratou-se da exibição de imagens presentes no trabalho de Snodgrass e Vanderwart (1980), que foram utilizadas de modo que o sujeito pudesse identificar uma condição de correspondência (BEGLEITER, 1999).

Figura 24 – Sistema 10-20 de posicionamento e escolha de canais.



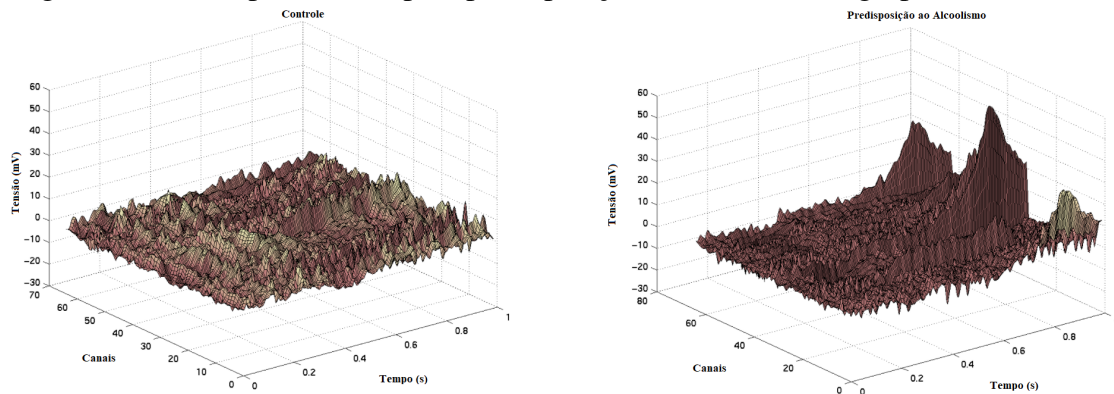
Fonte: Zhang *et al.* (1995). Modificado pelo Autor.

Dos 122 sujeitos analisados, 76 possuem predisposição ao alcoolismo e os demais são chamadas pelo autor de grupo de controle. Os estímulos visuais foram aplicados 120 vezes para cada pessoa, num período de um segundo, portanto, cada arquivo do paciente, contém 256 dados referentes um dos estímulos utilizados (BEGLEITER, 1999). A Figura 25 exhibe graficamente algumas diferenças entre os sinais de uma pessoa com a condição positiva, e a negativa. *A priori*, pode-se verificar que em termos de amplitude, os sinais do paciente com predisposição ao alcoolismo possui maior intensidade após o estímulo que o do grupo de controle.

5.3 Metodologia de Parametrização dos sinais

Nesta seção explica-se como é aplicada a parametrização do sinal EEG através de cada método utilizado no trabalho, seja clássico ou proposto. Esta explicação é realizada com-

Figura 25 – Exemplo de sinal para predisposição ao alcoolismo e grupo de controle.



Fonte: Begleiter (1999). Modificado pelo Autor.

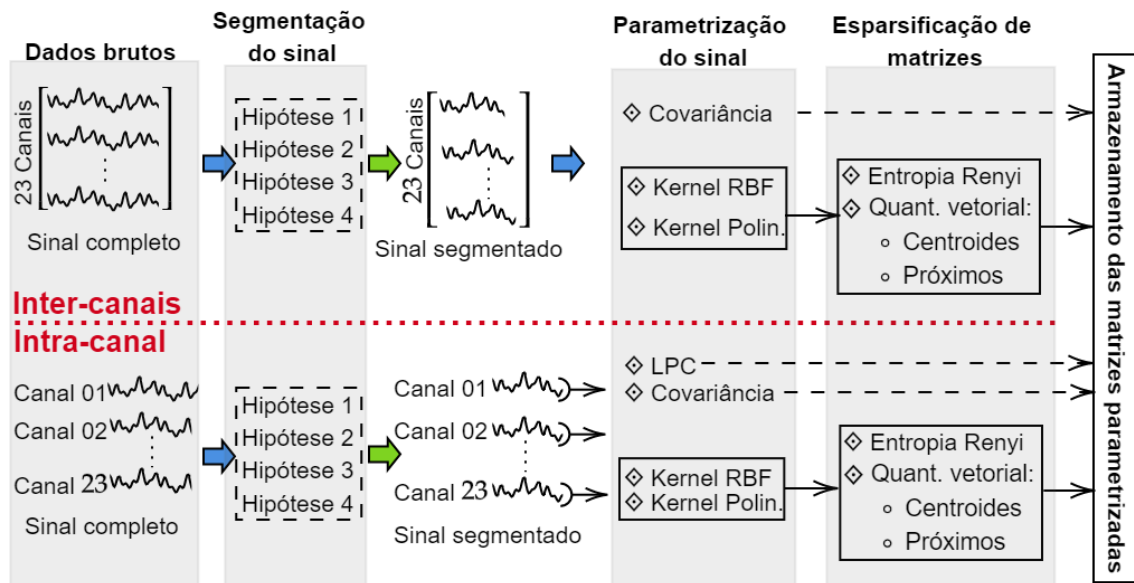
plementarmente ao que já foi descrito nos Capítulos 3 e 4 e foca-se na exibição de fluxogramas e algoritmos para entendimento do processo da aplicação dos métodos. Como os conjuntos de dados analisados possuem características particulares, faz-se necessário inicialmente uma explicação de como preparar seus arquivos para que em seguida sejam aplicadas as parametrizações. Em seguida, explica-se a metodologia de cada parametrização nas próximas subseções. Pode-se destacar ainda que todas as matrizes contendo os vetores de atributos produzidos por cada método de parametrização são armazenadas na memória do computador. Isto foi pensado de modo a reduzir o custo computacional na etapa de treinamento e teste dos classificadores.

O primeiro conjunto de dados dispõe de 130 arquivos com pelo menos um período convulsivo identificado. O processo para armazenamento das matrizes parametrizadas em apenas um desses arquivos, pode ser visualizado através do fluxograma exibido na Figura 26. Nesta é possível identificar que as abordagens intercanais e intracanal estão presentes e dífidas por uma linha tracejada horizontalmente. O processo inicia-se pela segmentação do sinal bruto através das hipóteses exibidas anteriormente pela Figura 23. Em seguida faz-se a aplicação dos métodos de parametrização, que podem ou não possuir uma etapa de esparsificação.

Destaca-se ainda que como a Figura 26 trata-se de uma explicação do processo de modo geral, deve-se abstrair a etapa das análises das épocas bem como os parâmetros de cada um dos métodos de parametrização e esparsificação. Então, as matrizes a serem armazenadas indicadas na última etapa da Figura, são compostas pelos vetores parametrizados de cada uma das épocas. Pode-se ainda explicar que na etapa de armazenamento, salvam-se 15 matrizes referentes aos possíveis métodos aplicados.

O processo de parametrização dos sinais presentes no segundo conjunto de dados,

Figura 26 – Fluxograma geral para armazenamento de matrizes parametrizadas no primeiro conjunto de dados.



Fonte: Elaborado pelo Autor.

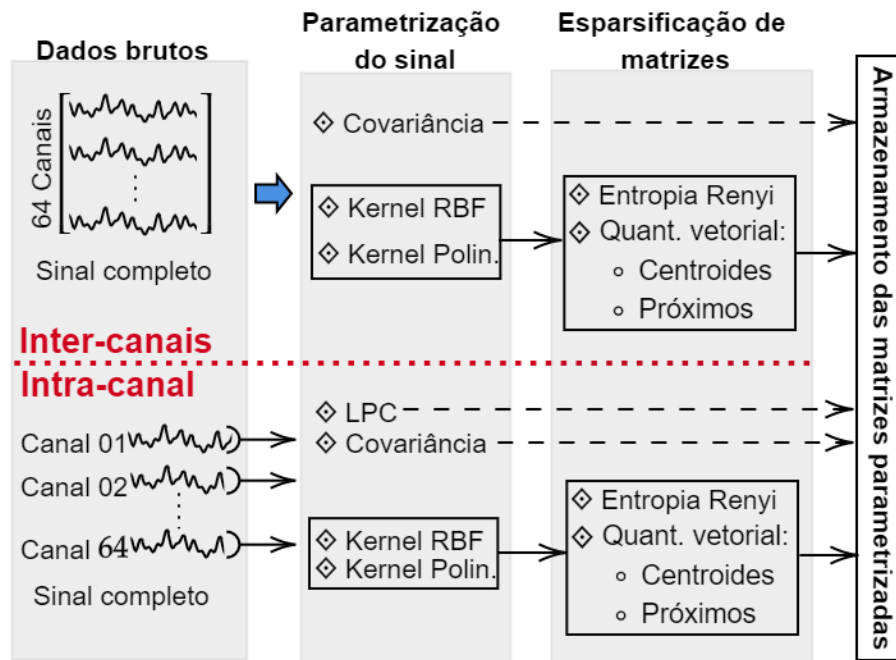
e sequencialmente o armazenamento das matrizes, é realizado conforme ilustrado pela Figura 27. Como dito anteriormente, cada arquivo disponível neste conjunto de dados possui apenas um segundo de dados (256 amostras), e por isso pensou-se em uma abordagem diferente do anterior, através do agrupamento dos dados de cada pessoa numa respectiva matriz. Como cada pessoa possui 120 arquivos, a sua matriz de dados \mathbf{X}_{subj} é composta pelos dados dos 64 canais e $120 \cdot (256)$ amostras. Portanto a Figura 27 explica a parametrização dos sinais de uma das pessoas, e não possui a etapa de segmentação do sinal.

Como dito anteriormente, na etapa de armazenamento são salvas 15 diferentes matrizes relacionadas aos métodos utilizados no presente trabalho. Essa alta quantidade é devida as variações dos métodos propostos. Estas variações podem ser visualizadas através da Figura 28, em que é possível verificar que a maioria das parametrizações propostas utilizadas no trabalho podem ser aplicadas nas duas abordagens, utilizam-se três métodos de esparsificação e as matrizes de kernel são compostas de duas formas diferentes. Além disto, utilizam-se os métodos clássicos de parametrização pela matriz de covariância (intercanais e intracanal) e LPC.

5.3.1 Metodologia de parametrização via LPC

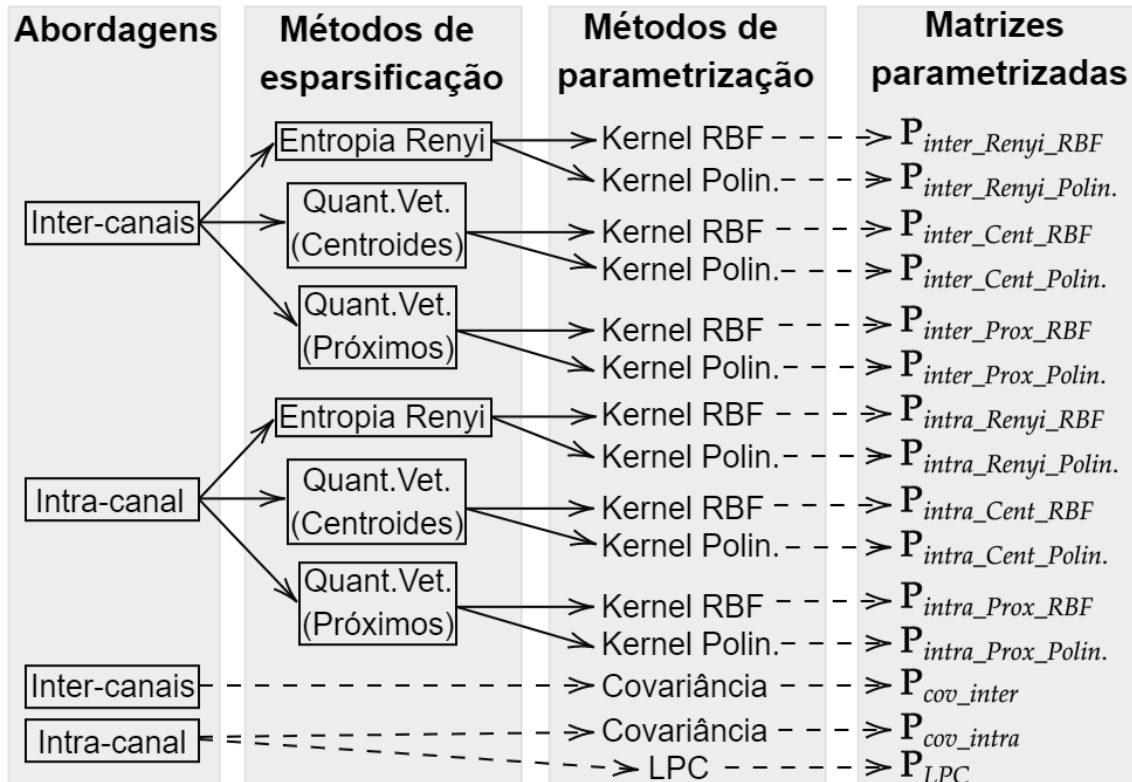
A metodologia de aplicação da parametrização pela estimação dos coeficientes LPC é exibida através do Algoritmo 3, que inicialmente necessita de uma massa de dados com ordem

Figura 27 – Fluxograma geral para armazenamento de matrizes parametrizadas no segundo conjunto de dados.



Fonte: Elaborado pelo Autor.

Figura 28 – Matrizes geradas no processo de parametrização a serem armazenadas.



Fonte: Elaborado pelo autor.

Algoritmo 3: Metodologia de parametrização via LPC.

Input: $\mathbf{X} \in \mathbb{R}^{K \times S}$, p
Output: Matriz parametrizada $\mathbf{P}_{LPC} \in \mathbb{R}^{[(K \cdot p) + 1] \times e}$
 Início;
 Inicializa a matriz \mathbf{P}_{cov} vazia;
 época $\leftarrow 1$;
while *Última época não é atingida*;
do
 $k \leftarrow 1$;
 Vetor \mathbf{v} é esvaziado;
 while $k \leq K$;
 do
 $\mathbf{s} \in \mathbb{R}^{1 \times N} \leftarrow \mathbf{X}$ [dados do canal k , dados da época atual];
 $\mathbf{aux} \in \mathbb{R}^{p \times 1} \leftarrow \text{LPC}(\mathbf{s}, p)$;
 concatena-se o vetor \mathbf{aux} em \mathbf{v} pela linha;
 $k \leftarrow k + 1$;
 end
 $\mathbf{v} \in \mathbb{R}^{[(K \cdot p) + 1] \times 1} \leftarrow$ adiciona-se o rótulo desta época no vetor \mathbf{v} ;
 concatena-se o vetor \mathbf{v} na matriz \mathbf{P} pela coluna;
 época \leftarrow época + 1;
end

$K \times S$ e a quantidade de coeficientes p . Destaca-se que para este e os demais métodos de parametrização, a massa de dados \mathbf{X} é a mesma, e pode ser vista como o sinal já segmentado de um paciente para o primeiro conjunto de dados, ou a massa de dados de um sujeito do segundo conjunto. Desta maneira, S representa a quantidade de amostras na massa de dados brutos. Esta se difere da quantidade de amostras N já apresentadas anteriormente, já que N é a quantidade de amostras em uma única época. Pode-se destacar ainda que a matriz resultante ao aplicar qualquer um dos métodos de parametrização, possui uma mesma quantidade e de colunas, referente a quantidade de épocas analisadas.

Por tratar de uma abordagem intracanal, pode-se verificar a presença de uma repetição que percorre cada canal. Dentro desta, aplica-se o método LPC ao sinal \mathbf{s} que resulta no vetor \mathbf{aux} com os p coeficientes. A cada sinal analisado por canal faz-se a concatenação destes vetores, compondo assim o vetor para uma época que possui ordem $K \cdot p \times 1$. Antes de realizar o avanço para uma próxima época, rotula-se este vetor de atributos com -1 caso a condição seja negativa ou $+1$ caso positiva. Ao final da época, faz-se o agrupamento do vetor construído na matriz \mathbf{P}_{LPC} .

5.3.2 Metodologia de parametrização pela matriz de covariância

Como descrito na seção 3.2, a parametrização pela matriz de covariância pode ser realizada pela abordagem intra e intracanal. Deste modo faz-se na presente subseção a exibição dos Algoritmos 4 e 5 que expressam o processo de parametrização em ambas abordagens em uma massa de dados \mathbf{X} . Destaca-se inicialmente que o Algoritmo 4 não necessita de algum parâmetro de entrada além dos dados. É possível ainda verificar que este Algoritmo possui apenas uma estrutura de repetição referente a cada época. Portanto, o método que estima a matriz de covariância e faz sua vetorização, recebe como argumento os dados \mathbf{D}^T , que contém os sinais presentes em todos os canais para uma única época.

Algoritmo 4: Metodologia de parametrização via matriz de Covariância (intercanais).

Input: $\mathbf{X} \in \mathbb{R}^{K \times S}$

Output: Matriz parametrizada $\mathbf{P}_{cov} \in \mathbb{R}^{K \cdot [(K+1)/2] + 1 \times e}$

Início;

Inicializa a matriz \mathbf{P} vazia;

época $\leftarrow 1$;

while Última época não é atingida;

do

$\mathbf{D} \in \mathbb{R}^{K \times N} \leftarrow \mathbf{X}[\text{dados dos } K \text{ canais, dados da época atual}]$;

$\mathbf{c} \in \mathbb{R}^{[K \cdot (K+1)]/2 \times 1} \leftarrow \text{VectCovar}(\mathbf{D}^T)$;

$\mathbf{c} \in \mathbb{R}^{[(K \cdot (K+1)/2) + 1] \times 1} \leftarrow$ adiciona-se o rótulo desta época no vetor \mathbf{c} ;

 concatena-se o vetor \mathbf{c} na matriz \mathbf{P} pela coluna;

 época \leftarrow época + 1;

end

O Algoritmo 5 no que lhe concerne, por se tratar de uma abordagem intracanal, utiliza-se da composição da matriz de Toeplitz \mathbf{R} através da Equação (3.7) fundamentada na subseção 3.2. A função que computa esta matriz necessita do parâmetro p que é similar ao utilizado no método LPC e define a quantidade de colunas em \mathbf{R} . Em sequência faz-se a aplicação do método de covariância e vetorização na matriz \mathbf{R} . Os processos subsequentes que finalizam o algoritmo e retorna a matriz com os sinais já parametrizados, são idênticos aos utilizados pela abordagem intracanal apresentada no Algoritmo 3.

5.3.3 Metodologia de parametrização através dos métodos propostos

A maneira como se aplicam os métodos propostos no presente trabalho, pode ser visualizado através dos Algoritmos 6 e 7. Identifica-se inicialmente que ambos necessitam que se

Algoritmo 5: Metodologia de parametrização via matriz de Covariância (Intracanal).

Input: $\mathbf{X} \in \mathbb{R}^{K \times S}$, p
Output: Matriz parametrizada $\mathbf{P}_{cov2} \in \mathbb{R}^{[K \cdot (p \cdot ((p+1)/2)) + 1] \times e}$
 Início;
 Inicializa a matriz \mathbf{P}_{cov2} vazia;
 época $\leftarrow 1$;
while Última época não é atingida;
do
 $k \leftarrow 1$;
 Vetor \mathbf{v} é esvaziado;
 while $k \leq K$;
 do
 $\mathbf{s} \in \mathbb{R}^{1 \times N} \leftarrow \mathbf{X}[k, \text{dados da época atual}]$;
 $\mathbf{R} \in \mathbb{R}^{(N-p+1) \times p} \leftarrow \text{Toeplitz}(\mathbf{s}, p)$;
 $\mathbf{c} \in \mathbb{R}^{[p \cdot (p+1)/2] \times 1} \leftarrow \text{VectCovar}(\mathbf{R})$;
 concatena-se o vetor \mathbf{c} em \mathbf{v} pela linha;
 $k \leftarrow k + 1$;
 end
 $\mathbf{v} \in \mathbb{R}^{[K \cdot (p \cdot (p+1)/2) + 1] \times 1} \leftarrow$ adiciona-se o rótulo desta época no vetor \mathbf{v} ;
 concatena-se o vetor \mathbf{v} na matriz \mathbf{P}_{cov2} pela coluna;
 época \leftarrow época + 1;
end

defina qual função de kernel será utilizada, bem como qual método de esparsificação é aplicado na massa de dados. Independente qual seja a função de kernel escolhida, está normalmente necessita da definição de valores de hiperparâmetros. Desta maneira, destaca-se que a escolha destes é abordada na próxima seção. Todavia, o método de esparsificação possui um parâmetro livre M que define a quantidade de vetores relevantes a serem agrupados na matriz \mathbf{E} . Como dito seção 4.2, dois métodos de esparsificação foram escolhidos para o presente trabalho, portanto, faz-se necessária a explicação que $M = k$ ao utilizar o método k -Médias e $M = m$ caso se utilize o critério de esparsificação pela entropia de Renyi.

O Algoritmo 6 exhibe a aplicação dos métodos propostos através da abordagem intercanais. Logo, a cada época tem-se uma matriz com os dados presentes em todos os canais. Nesta matriz faz-se a seleção do dicionário de vetores relevantes ao utilizar o método de esparsificação escolhido, compondo assim a matriz \mathbf{E} . Em sequência aplica-se a função de kernel, que recebe \mathbf{E} como argumento, bem como seus parâmetros necessários. Destaca-se que a função *KernelVect* realiza tanto a estimação como a vetorização da matriz de kernel que tem ordem $M \times M$, formando assim o vetor \mathbf{v} . Este por sua vez é rotulado como feito para os métodos anteriores. O processo é repetido até que a última época seja atingida e a matriz \mathbf{P}_{kernel} , seja

Algoritmo 6: Metodologia de parametrização via matriz de Kernel (Intercanais).

Input: $\mathbf{X} \in \mathbb{R}^{K \times S}$, Função de Kernel (RBF ou Polinomial), Método de Esparsificação e valor de M

Output: Matriz parametrizada $\mathbf{P}_{kernel-inter} \in \mathbb{R}^{[(M \cdot (M+1))/2] + 1 \times e}$

Início;

Inicializa a matriz $\mathbf{P}_{kernel-inter}$ vazia;

época $\leftarrow 1$;

while Última época não é atingida;

do

$\mathbf{D} \in \mathbb{R}^{K \times N} \leftarrow \mathbf{X}$ [dados dos K canais, dados da época atual];

$\mathbf{E} \in \mathbb{R}^{M \times K} \leftarrow \text{Esparsificação}(\mathbf{D}^T, M)$;

$\mathbf{v} \in \mathbb{R}^{[(M \cdot (M+1))/2] \times 1} \leftarrow \text{KernelVect}(\mathbf{E}, \text{Parâmetros Kernel})$;

$\mathbf{v} \in \mathbb{R}^{[(M \cdot (M+1))/2] + 1 \times 1} \leftarrow$ adiciona-se o rótulo desta época no vetor \mathbf{v} ;

 concatena-se o vetor \mathbf{v} na matriz $\mathbf{P}_{kernel-inter}$ pela coluna;

 época \leftarrow época + 1;

end

completamente construída.

A utilização dos métodos propostos pela abordagem intracanal pode ser vista através do Algoritmo 6, que além de necessitar dos mesmos parâmetros de entrada da versão intercanais, necessita da definição do valor de p . Portanto, da mesma maneira como ocorre para o método de covariância intracanal, utiliza-se da função que compõe a matriz de Toeplitz \mathbf{R} a partir do sinal s referente ao canal k analisado. A abordagem proposta no presente trabalho faz a esparsificação da matriz \mathbf{R} afim de a compor o dicionário de vetores relevantes \mathbf{E} . De posse deste, aplica-se a função de kernel que estima uma matriz de ordem $M \times M$ e faz sua vetorização e normalização. Por tratar da abordagem intracanal, este procedimento é repetido concatenando os vetores pela linha para cada canal. Após este procedimento, faz-se a rotulação do vetor de atributos e para este ser agrupado aos demais vetores de cada época na matriz $\mathbf{P}_{Kernel-intra}$.

5.4 Metodologia para treino e teste dos classificadores.

Com a explicação de como a parametrização é realizada e suas matrizes armazenadas na memória do computador, faz-se necessário descrever como os modelos gerados pelos algoritmos de classificação são estimados e executados. A explicação desta etapa é necessária, pois a maneira de como a mesma é aplicada difere-se para os dois conjuntos de dados. Como dito anteriormente, por existirem algumas variações dos métodos propostos, o processo de parametrização em uma massa de dados brutos do sinal EEG, gera 15 matrizes. Portanto, cada um

Algoritmo 7: Metodologia de parametrização via matriz de Kernel (Intracanal).

Input: $\mathbf{X} \in \mathbb{R}^{K \times S}$, Função de Kernel (RBF ou Polinomial), Método de Esparsificação (Entropia Reniy ou Quantização vetorial), M, p

Output: Matriz parametrizada $\mathbf{P}_{Kernel-intra} \in \mathbb{R}^{[K \cdot (M \cdot (M+1)/2) + 1] \times e}$

Início;

Inicializa a matriz $\mathbf{P}_{Kernel-intra}$ vazia;

época $\leftarrow 1$;

while Última época não é atingida;

do

$k \leftarrow 1$;

 Vetor \mathbf{v} é esvaziado;

while $k \leq K$;

do

$\mathbf{s} \in \mathbb{R}^{1 \times N} \leftarrow \mathbf{X}[\text{canal } k, \text{ dados da época atual}]$;

$\mathbf{R} \in \mathbb{R}^{(N-p+1) \times p} \leftarrow \text{MatRegressores}(\mathbf{s}, p)$;

$\mathbf{E} \in \mathbb{R}^{M \times p} \leftarrow \text{Esparsificação}(\mathbf{R})$;

$\mathbf{aux} \in \mathbb{R}^{[M \cdot (M+1)/2] \times 1} \leftarrow \text{FunçãoKernelVect}(\mathbf{E})$;

 concatena-se o vetor \mathbf{aux} em \mathbf{v} pela linha;

$k \leftarrow k + 1$;

end

$\mathbf{v} \in \mathbb{R}^{[K \cdot (M \cdot (M+1)/2) + 1] \times 1} \leftarrow$ adiciona-se o rótulo desta época no vetor \mathbf{v} ;

 concatena-se o vetor \mathbf{v} na matriz $\mathbf{P}_{Kernel-intra}$ pela coluna;

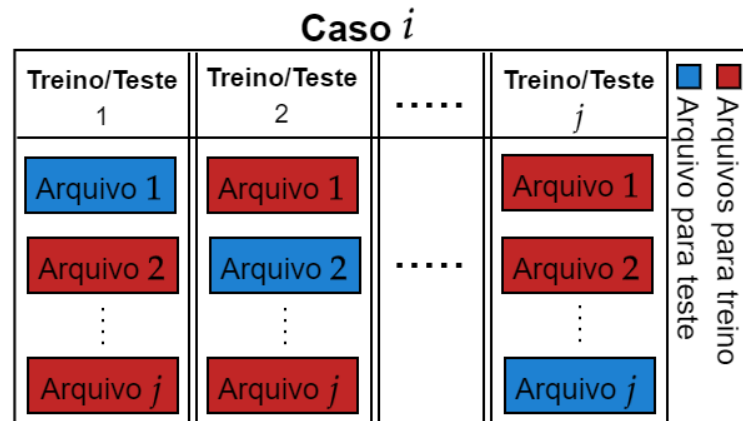
 época \leftarrow época + 1;

end

dos dois classificadores utilizados no presente trabalho, faz *a priori*, a estimação de 15 diferentes modelos.

O procedimento para treino e teste dos classificadores no primeiro conjunto de dados, foi pensado de maneira a simular um sistema que faz a identificação de sinais epiléticos de um paciente em tempo real, a partir do modelo gerado de sinais armazenados previamente. Como se trata apenas de uma simulação, realizou-se um processo de validação cruzada em cada caso do conjunto de dados, como pode ser visualizado através da Figura 29. Nesta é possível verificar que se utiliza da abordagem em dobras, ou seja, para cada caso, o treinamento e teste do modelo são repetidos j vezes. Esta etapa no que lhe concerne, utiliza-se de um arquivo para teste e os demais para treino.

A etapa de treinamento utiliza-se das matrizes com os sinais já parametrizados armazenados previamente. Como o processo de parametrização em um arquivo gera 15 matrizes, estas são agrupadas com as 15 geradas por cada arquivo do conjunto de treinamento. Desta maneira o algoritmo de classificação escolhido, faz a estimação de 15 diferentes modelos (um

Figura 29 – Processo de validação cruzada em j dobras.

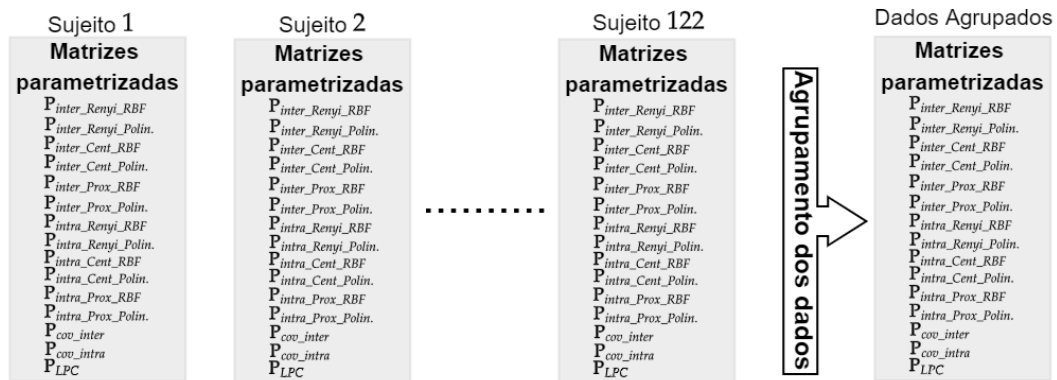
Fonte: Elaborado pelo autor.

para cada método de parametrização) em sua etapa de treinamento. A etapa de teste é realizada de maneira diferente, pois, como já dito pensou-se numa simulação de um sistema que faz a classificação dos sinais nas épocas em convulsivos ou não. Portanto, o arquivo utilizado como teste é percorrido do início ao fim sem alguma aplicação de segmentação. Todavia, manteve-se ainda a metodologia de realizar a parametrização no curto intervalo de tempo chamado de época. Deste modo, a cada época realiza-se a parametrização gerando um vetor de atributos, no que lhe concerne é utilizado como argumento no processo de predição do classificador.

O procedimento de treino e teste para os sinais presentes no segundo conjunto de dados possui similaridades e diferenças com relação ao primeiro conjunto. Isto porque o problema de classificação abordado trata-se da distinção entre sinais de pessoas que possuem ou não uma predisposição genética ao alcoolismo. Portanto, o modelo gerado pela etapa de treinamento e teste do classificador, utiliza dados de diferentes sujeitos. Como dito anteriormente, cada sujeito possui armazenado na memória, 15 matrizes referentes ao processo de parametrização. Realizou-se então o agrupamento dessas matrizes para todos os sujeitos, formando assim um único grupo de 15 matrizes como pode ser visualizado através da Figura 30.

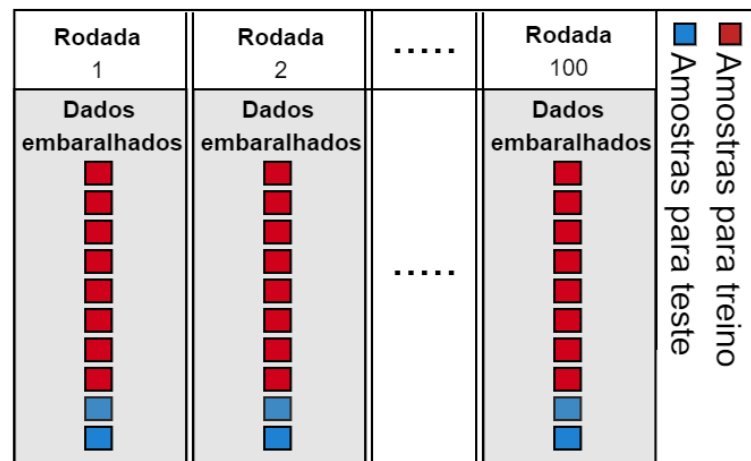
As matrizes desse grupo final são utilizadas para treinamento e teste, avaliando assim a capacidade de classificação para cada método de parametrização aplicado. Na etapa de treinamento e teste, utilizou-se a divisão dos dados em 80% para treino e 20% para teste. A fim de verificar uma maior variabilidade do conjunto de dados, aplicou-se a técnica ilustrada através da Figura 31. Nesta é possível verificar que a etapa de treino e teste é repetida 100 vezes. Além disto, a cada rodada, o conjunto de dados é embaralhado utilizando uma função de permutação aleatória e em seguida faz-se a divisão do conjunto de treino e teste.

Figura 30 – Agrupamento dos grupos de matrizes parametrizadas do segundo conjunto de dados.



Fonte: Elaborado pelo autor.

Figura 31 – Processo de treino e teste para segundo conjunto de dados.



Fonte: Elaborado pelo autor.

5.5 Avaliação de desempenho

Como dito anteriormente, utiliza-se de um conjunto de treinamento para construir um modelo que faz predições a partir de novas amostras. Contudo, é necessário verificar o desempenho destes modelos, ou seja, utilizar de métricas ou gráficos que avaliem a capacidade de um modelo realizar suas predições corretamente. A partir da quantidade de classes, pode-se dizer que existem dois tipos de problemas de classificação: binária (com apenas duas classes); múltiplas classes (acima de duas classes) (THARWAT, 2020). Para o presente trabalho os dois conjuntos de dados utilizados são enquadrados no problema de classificação binário. Portanto, as classes assumem valores +1 para condição positiva ou -1 para condição negativa.

Uma análise de desempenho comumente utilizada é através da matriz de confusão.

Esta matriz de ordem 2×2 pode ser composta como ilustrado pela Figura 32 que contém as informações das quantidades de Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Negativo (FN) e Falso Positivo (FP). Os valores de VP e VN representam respectivamente a quantidade de predições corretas para a condição positiva e negativa. FP é a quantidade de predições erradas para a condição atual negativa e FN é a soma de predições realizadas de forma errada para a condição atual positiva (MCGWIRE; FISHER, 2001).

Figura 32 – Exemplo de Matriz de Confusão.

		Real	
		Condição Positiva	Condição Negativa
Predito	Condição Positiva	Verdadeiro Positivo [1,1]	Falso Positivo [1,-1]
	Condição Negativa	Falso Negativo [-1,1]	Verdadeiro Negativo [-1,-1]

Fonte: Elaborado pelo autor.

Com a composição da matriz de confusão, pode-se calcular as métricas de desempenho do classificador. Uma das mais utilizadas é a acurácia que expressa a razão entre as quantidades preditas corretamente e todas as quantidades como definida na seguinte equação:

$$Acc = \frac{VP + VN}{VP + VN + FP + FN}. \quad (5.1)$$

Esta medida também é conhecida como taxa de acerto e seu complemento $1 - Acc$ mede a taxa de erro (SOKOLOVA *et al.*, 2006).

Como a acurácia é sensível para problemas em que se tenha uma quantidade desbalanceadas de classes, faz-se o uso de duas métricas que expressam a capacidade de classificação negativa ou positiva. A primeira chamada de sensibilidade, expressa através da equação:

$$Sensibilidade = \frac{VP}{VP + FN}, \quad (5.2)$$

que representa a razão entre a quantidade de predições corretas positivas e todas as predições positivas realizadas. Ao passo que a especificidade é a proporção de classificação correta da

condição negativa por todas as predições da condição negativa como definida na seguinte equação (MCGWIRE; FISHER, 2001; SOKOLOVA *et al.*, 2006):

$$Especificidade = \frac{VN}{VN + FP}. \quad (5.3)$$

Como o problema em um dos conjuntos de dados possui uma quantidade desbalanceada de classes, fez-se o uso de uma métrica que incorpora a acurácia, sensibilidade e especificidade em um só valor. Tal medida envolve o cálculo da média geométrica entre as três taxas, sendo definida como

$$Média\ Geométrica = \sqrt[3]{Acurácia \cdot Sensibilidade \cdot Especificidade} \quad (5.4)$$

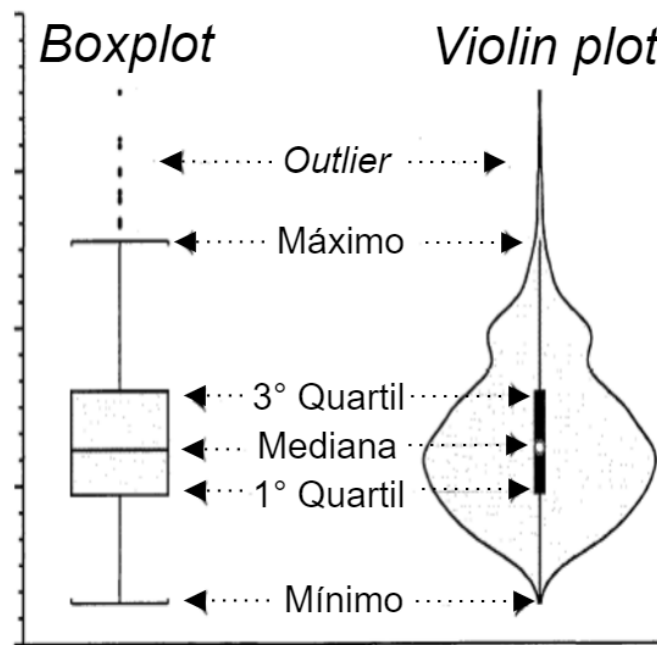
Para exibir o desempenho do classificador graficamente, escolheu-se a utilização de *violin plot*. Esta ferramenta é uma derivação do diagrama de caixas (*boxplot*) que exibe graficamente grupos de dados numéricos através de seus quartis. O *boxplot* como pode ser visualizado na Figura 33 é composto por cinco elementos:

- **Mínimo:** menor valor entre os dados, excluindo um possível *outlier*;
- **Máximo:** maior valor entre os dados, excluindo um possível *outlier*;
- **Mediana:** calculada de todos os pontos do conjunto de dados;
- **Primeiro Quartil:** representando a mediana da metade inferior dos dados;
- **Terceiro Quartil:** representando a mediana da metade superior dos dados.

A fim de expressar mais informações sobre a distribuição dos dados, o *violin plot* exibido através da Figura 33, além de apresentar as informações contidas no *box plot*, mostra a densidade de probabilidade que é comumente suavizada através de um estimador de densidade kernel (HYNDMAN; SHANG, 2010).

5.6 Considerações Finais

No presente capítulo foram descritas as características dos conjuntos de dados escolhidos para avaliar o desempenho das parametrizações do sinal EEG. Em seguida, fez-se uma explicação sobre a metodologia escolhida para aplicação de cada uma das parametrizações utilizadas, através de algoritmos e fluxogramas. Como foram escolhidos dois conjuntos de dados diferentes, explicam-se como os dados parametrizados, em cada um destes casos, são divididos em subconjuntos de treino e teste. O capítulo se encerra pela descrição das métricas utilizadas para avaliar os desempenhos de cada modelo gerado ao longo do trabalho.

Figura 33 – Comparação entre *Boxplot* e *Violin plot*.

Fonte: Hu (2020). Modificado pelo Autor.

6 RESULTADOS E DISCUSSÕES

No presente capítulo serão descritos os resultados e as discussões obtidas ao aplicar os métodos propostos nos dois conjuntos de dados. Para cada conjunto, faz-se inicialmente a apresentação dos resultados referentes ao processo de exploração dos hiperparâmetros. Após a definição destes, exibem-se os resultados com os comparativos entre métodos propostos e clássicos. Vale ressaltar que pelo fato do primeiro conjunto de dados possuir diversos pacientes (podendo ser entendidos cada um como um subconjunto), apenas alguns foram escolhidos para elaborar os resultados no presente capítulo. Essa escolha foi realizada de modo a exibir os resultados mais representativos das seguintes possibilidades: resultados inferiores aos métodos clássicos; resultados superiores aos métodos clássicos; resultados similares aos métodos clássicos. Os resultados dos outros casos, são exibidos no Apêndice F, através os gráficos de *violin plot* para as taxas de sensibilidade, acurácia e especificidade. No final do presente capítulo, são apresentados as análises de custo computacional de cada método aplicado no trabalho.

Por fim, se faz necessária uma explicação das abreviações utilizadas nas figuras ao longo do capítulo. Como dito anteriormente, são avaliados no presente trabalho 12 variações dos métodos propostos e três métodos clássicos. Para realizar a identificação dos métodos nos gráficos presentes no capítulo atual, foi necessário abreviar cada uma das nomenclaturas. Tais abreviações são exibidas através do Tabela 2.

Tabela 2 – Nomenclaturas utilizadas nos gráficos dos resultados.

Abreviação	Explicação do método de parametrização
kRBF-C	Método de esparsificação pelo k -Médias utilizando os k vetores protótipos em conjunto com a matriz de kernel RBF
kRBF-P	Método de esparsificação pelo k -Médias utilizando os k vetores mais próximos aos protótipos em conjunto com a matriz de kernel RBF
kPoly-C	Método de esparsificação pelo k -Médias utilizando os k vetores protótipos em conjunto com a matriz de kernel polinomial
kPoly-P	Método de esparsificação pelo k -Médias utilizando os k vetores mais próximos aos protótipos em conjunto com a matriz de kernel polinomial
mRBF	Método de esparsificação pelo critério de maximização da entropia de Renyi em conjunto com a matriz de kernel RBF
mPoly	Método de esparsificação pelo critério de maximização da entropia de Renyi em conjunto com a matriz de kernel polinomial

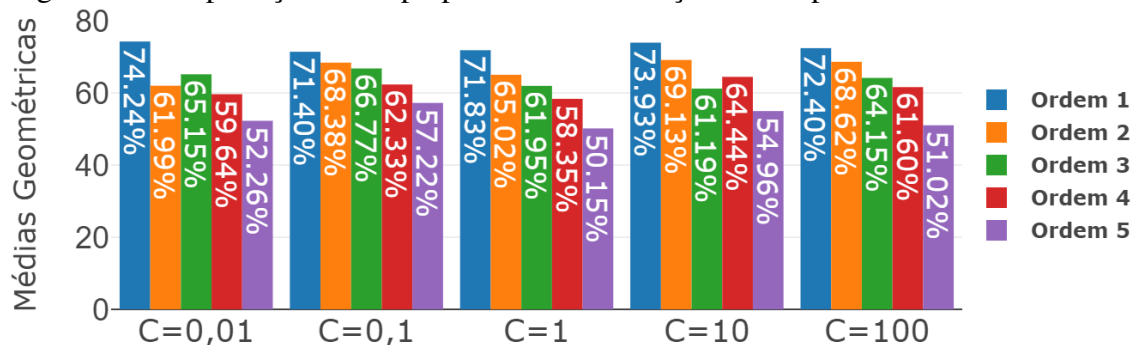
Fonte: Elaborado pelo Autor.

6.1 Resultados para o primeiro conjunto de dados

Como dito anteriormente, antes da aplicação dos métodos nos arquivos dos pacientes escolhidos, fez-se uma exploração dos hiperparâmetros de cada um dos métodos. É possível destacar que esse levantamento foi realizado utilizando apenas os arquivos do caso **CHB01**, pela hipótese de segmentação do sinal de número 04 e utilizando o classificador OLS. Ressalta-se ainda que todas as métricas exibidas nesta subseção são referentes a média geométrica obtida pela Equação (5.4).

Inicialmente a exploração foi realizada na função de kernel polinomial, ao variar a ordem do polinômio com os valores $d = \{1, 2, 3, 4, 5\}$ e o parâmetro $C = \{0,01, 0,1, 1, 10, 100\}$. Destaca-se que essa exploração foi realizada utilizando apenas o modo intracanal ($p = 9$) e esparsificação pelo método k -Médias ($k = 5$) ao escolher apenas os k vetores mais próximos dos protótipos. Nesta exploração, foi possível compor o gráfico da Figura 34, em que se identificou que ao aumentar a ordem do polinômio, os valores de médias geométricas tem característica de queda. Destaca-se ainda ao aplicar o método com $C = 0,01$, foi possível obter os maiores resultados em cada ordem polinomial, portanto, os parâmetros $d = 1$ e $C = 0,01$ foram definidos.

Figura 34 – Exploração dos hiperparâmetros da função kernel polinomial.

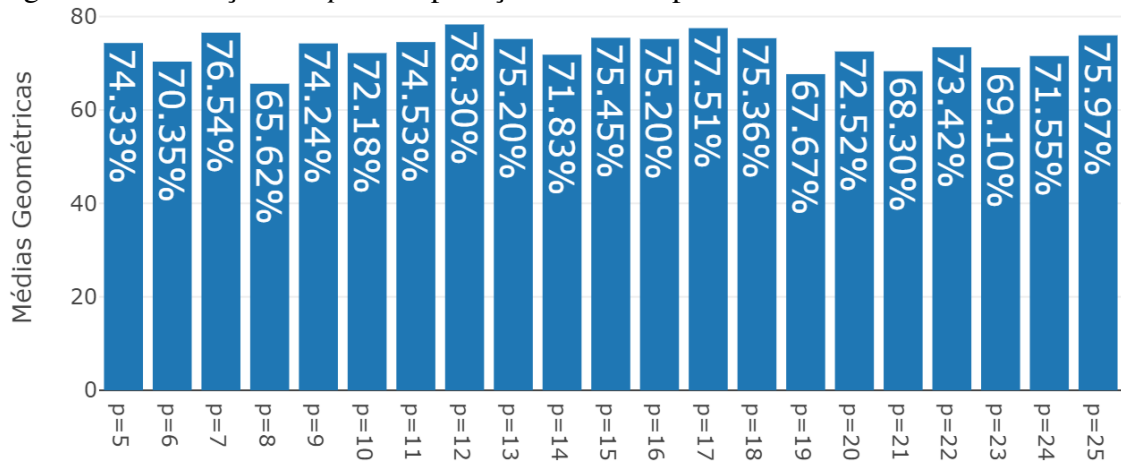


Fonte: Elaborado pelo Autor.

Em sequência, buscou-se realizar uma variação de p para identificar um possível valor que aumentasse a média geométrica obtida. Esta variação foi feita de modo a verificar um padrão de queda nos valores, com o crescimento de p . A Figura 35 representa os resultados obtidos nesta variação e destaca-se que foi definido o valor de $p = 12$ para a utilização do kernel polinomial. Pode-se verificar ainda que com o aumento de p a característica de queda não foi identificada e sim uma variação em torno do valor de $72,74\% \pm 3,24\%$

A exploração realizada em sequência foi pensada diferentemente, para que fosse

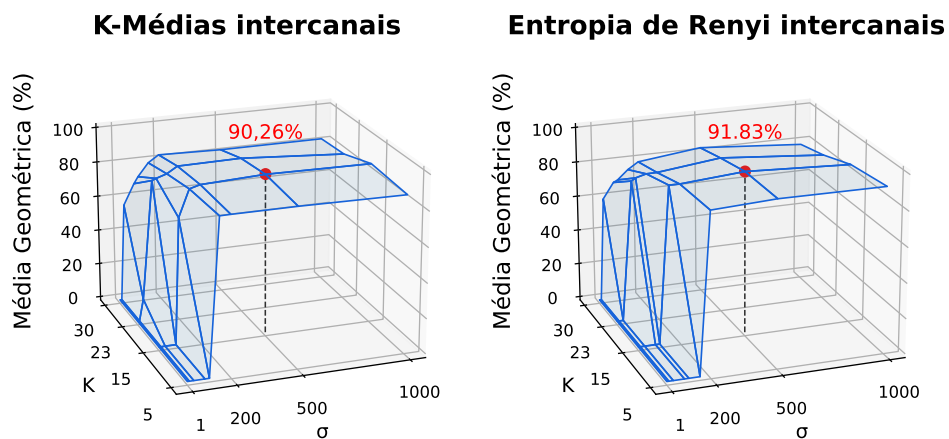
Figura 35 – Variações de p com aplicação no kernel polinomial.



Fonte: Elaborado pelo Autor.

possível abranger uma maior gama de hiperparâmetros. Dessa maneira, buscou-se avaliar a abordagem intercanais, a intracanal e os parâmetros da função de kernel RBF. Sendo assim, os parâmetros de σ pertencente ao kernel RBF elencados para teste foram $\{1, 10, 30, 80, 100, 150, 200, 500, 1000\}$ e para a abordagem intercanais os valores $k = m = \{5, 15, 23, 30\}$. Com isto, foi possível compor os gráficos presentes na Figura 36, que exibe as explorações realizadas tanto para o método k -Médias como pelo critério de entropia de Renyi. Pode-se verificar que em ambos gráficos os valores de $\sigma = 500$ e $k = m = 23$ fazem com que a média geométrica seja superior as demais configurações.

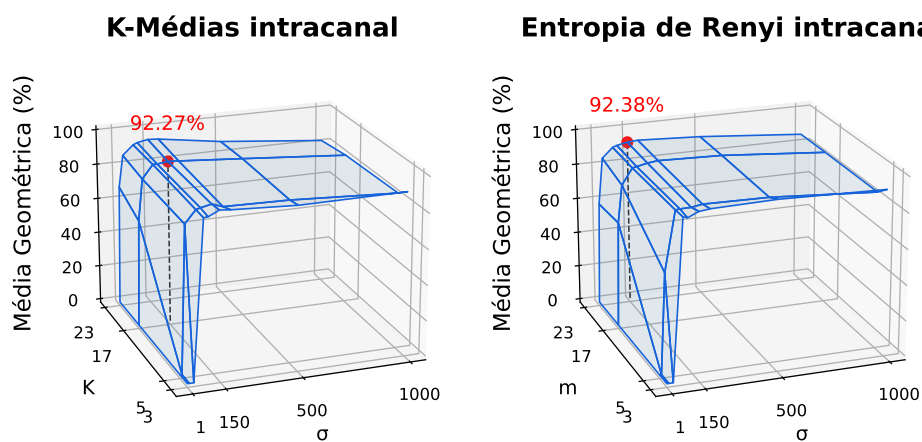
Figura 36 – Exploração kernel RBF pela abordagem intercanais.



Fonte: Elaborado pelo Autor.

Para a abordagem intracanal, escolheram-se os valores de $k = m = \{3, 5, 17, 23\}$ e os valores do parâmetro σ foram mantidos. Dessa maneira, após a realização da exploração, foi feita a composição dos gráficos na Figura 37, que ilustra as explorações dos métodos de esparsificação k -Médias e do critério de entropia de Renyi. Verifica-se através dos gráficos, que ao utilizar os parâmetros $\sigma = 150$, $k = 17$ e $m = 23$ as médias geométricas encontradas são superiores as outras variações de hiperparâmetros.

Figura 37 – Exploração kernel RBF e abordagem intracanal.

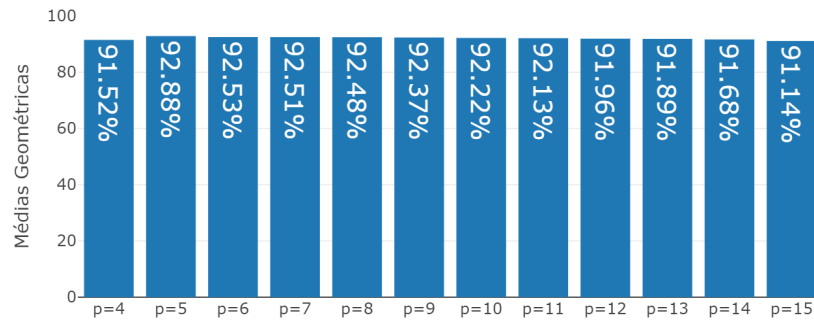


Fonte: Elaborado pelo Autor.

Por fim, realizou-se uma exploração do parâmetro p com os valores presentes na Figura 38. Nesta é possível identificar que ao aumentar a magnitude de p , ocorre um leve decréscimo dos valores de média geométrica. Com a obtenção deste resultado, definiu-se o valor de $p = 5$ para as abordagens intracanal que utilizam a função de kernel RBF.

Após a definição de hiperparâmetros, foram escolhidos quatro casos do conjunto de dados e fez-se a aplicação dos métodos propostos no presente trabalho. A escolha destes casos foi realizada pensando em discutir resultados relevantes obtidos ao se utilizar o conjunto total. Vale ressaltar que na etapa de classificação, o método OLS e o CLG, foram utilizados para compor os resultados. Além disto, pode-se perceber ao decorrer desta seção, que alguns dos métodos de parametrização são omitidos para que haja um destaque naqueles que obtiveram um melhor resultado. Para cada um dos casos exibidos, são realizadas análises comparativas entre os desempenhos obtidos pelos métodos clássicos e propostos, porém, uma análise mais detalhada é realizada para a métrica de sensibilidade.

Figura 38 – Exploração de p em modo intracanal para kernel RBF com esparsificação pela entropia de Renyi.



Fonte: Elaborado pelo Autor.

6.1.1 Resultados caso CHB01

O caso **CHB01** possui sete arquivos com crises epiléticas identificadas, suas aquisições completas tem duração de uma hora e a quantidade média de períodos convulsivos é de 56,83s. Inicialmente, para a hipótese de segmentação 01, são exibidas as médias geométricas obtidas ao aplicar os dois classificadores, através da Figura 39. Nesta Figura é possível verificar que a análise pelo método de covariância na abordagem intercanais, possui um melhor resultado que os métodos propostos desta mesma abordagem. Porém, ao analisar estes resultados para o classificador linear gaussiano, é possível verificar que o desempenho pela abordagem intercanais é superior. Nesta diferença, foi possível verificar que os métodos em que se utilizou a matriz de kernel RBF obtiveram resultados próximos ao da matriz de covariância, porém, ainda inferiores.

Outra análise que pode ser realizada na Figura 39, é que pela abordagem intracanal os desempenhos são bem diferentes, já que as médias geométricas obtidas pelos métodos que se utilizam da matriz de kernel RBF são superiores aos demais, independentemente da abordagem aplicada. Este ocorrido foi observado tanto para a aplicação do OLS como no CLG. Porém, pode-se verificar um desempenho inferior nos exemplos em que se utilizou a matriz de kernel polinomial.

Para analisar as sensibilidades obtidas pelos classificadores, compôs-se o *violin plot* exibido pela Figura 40. Nesta é possível verificar que ao utilizar a matriz de kernel polinomial pela abordagem intercanais, a capacidade de identificar corretamente as crises epiléticas, é inferior aos demais métodos. Além disto, verifica-se pelos dois classificadores que as sensibilidades são superiores e com menor dispersão, nos exemplos que utilizaram a matriz de kernel RBF. Tratando dos métodos clássicos, percebe-se que a matriz de covariância intercanais e o método

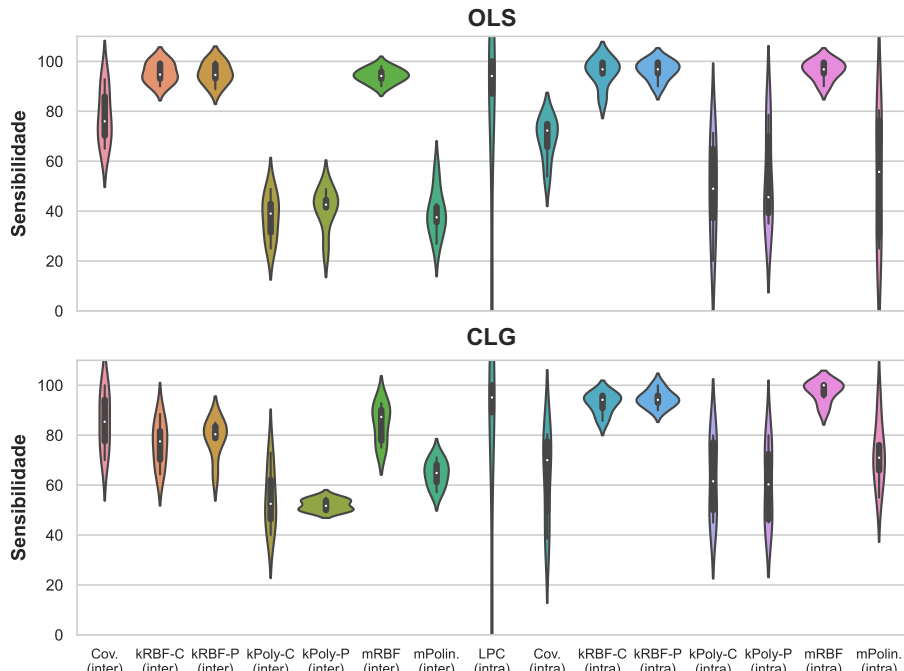
Figura 39 – **CHB01**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.



Fonte: Elaborado pelo Autor.

LPC, se aproximam dos melhores métodos mesmo que o LPC tenha tido uma alta dispersão.

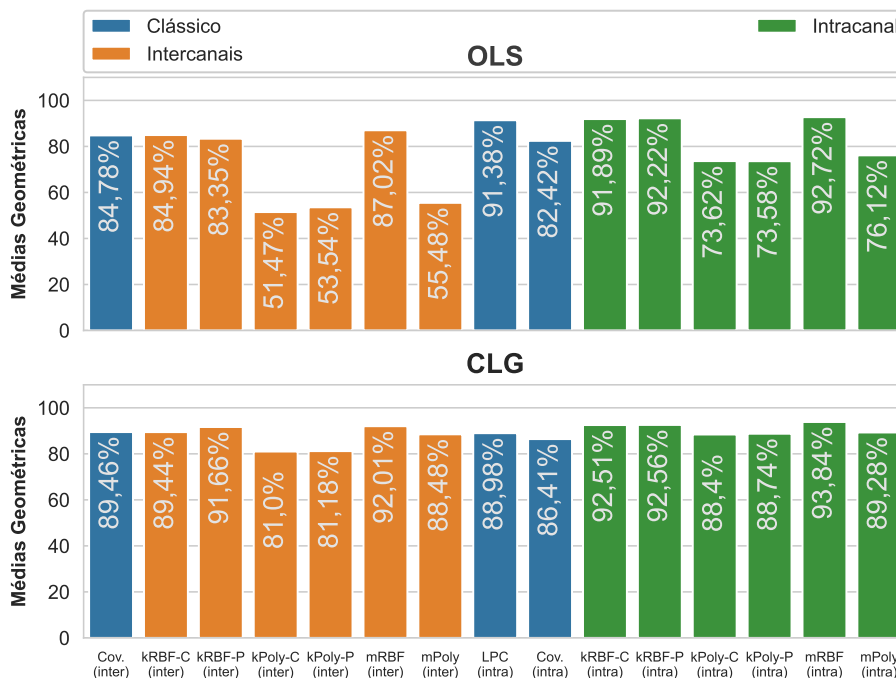
Figura 40 – **CHB01**: Comparativo das sensibilidades entre OLS e CLG para segmentação 01.



Fonte: Elaborado pelo Autor.

Ao analisar os resultados obtidos pela hipótese 02 de segmentação do sinal, pode-se compor os mesmos gráficos exibidos anteriormente. As médias geométricas nesta configuração são exibidas pela Figura 41, em que é perceptível um aumento no desempenho ao comparar com a segmentação anterior e nos métodos que se utilizou a matriz de kernel RBF. Percebe-se ainda que pela abordagem intracanal, este kernel propicia ao método proposto uma medida superior aos métodos clássicos, mesmo que o LPC tenha atingido valor próximo. Além disto, as características de menores desempenhos ainda são constatadas pelos métodos que fizeram o uso da matriz de kernel polinomial.

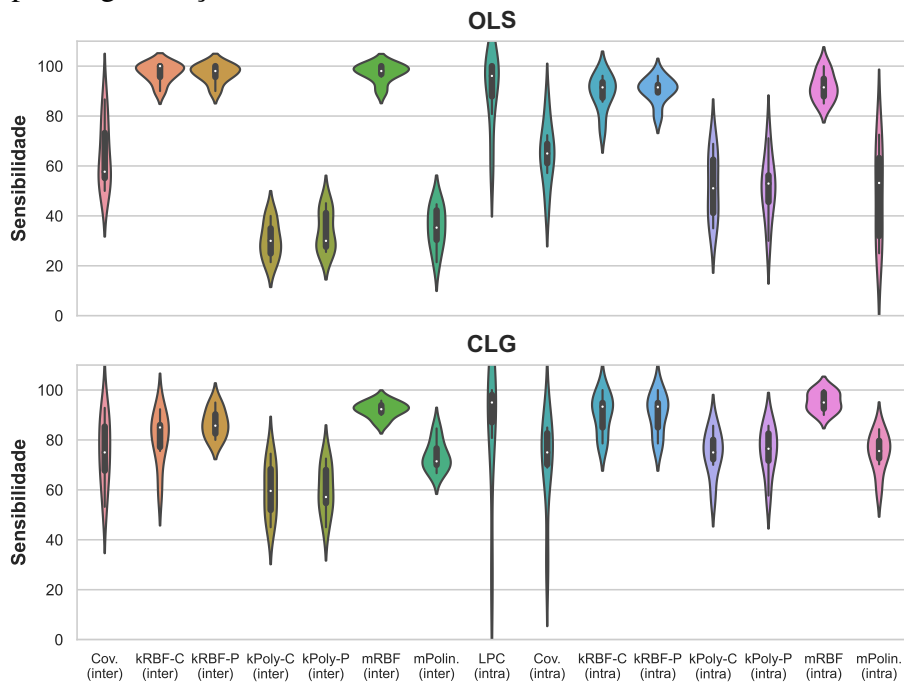
Figura 41 – **CHB01**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.



Fonte: Elaborado pelo Autor.

A análise das sensibilidades realizada para a segmentação 02, pode ser visualizada através da Figura 42. Percebe-se inicialmente através desta, que para ambos os classificadores, os valores obtidos ao aplicar o método da covariância intercanais, é inferior ao utilizar a segmentação 01. A Figura ainda exhibe que as sensibilidades obtidas pelos métodos que usam a matriz de kernel RBF, são altas e menos dispersas que os outros métodos. Porém, ao analisar essas taxas obtidas no CLG, percebe-se uma maior dispersão.

Figura 42 – **CHB01**: Comparativo das sensibilidades entre OLS e CLG para segmentação 02.

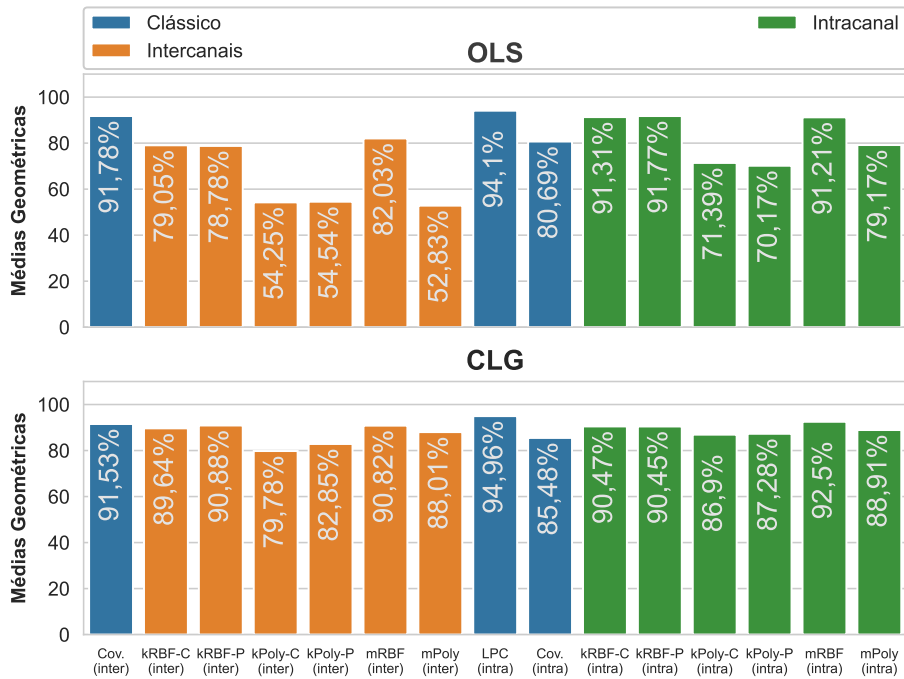


Fonte: Elaborado pelo Autor.

Fazendo o uso da hipótese de segmentação 03, pode-se verificar através da Figura 43, que o padrão obtido é diferente ao observado anteriormente. Isto porque o método LPC para esta hipótese possui média geométrica superior aos demais métodos, com valor de 94,1% para OLS e 94,96% para o CLG. Mesmo com tal ocorrido, os métodos que utilizam a matriz de kernel ainda puderam ter médias geométricas superiores às obtidas pela matriz de covariância. Ao comparar os resultados entre os classificadores, é possível verificar que no CLG há um aumento no desempenho dos métodos pela abordagem intercanais, porém, a análise da sensibilidade ainda precisa ser realizada.

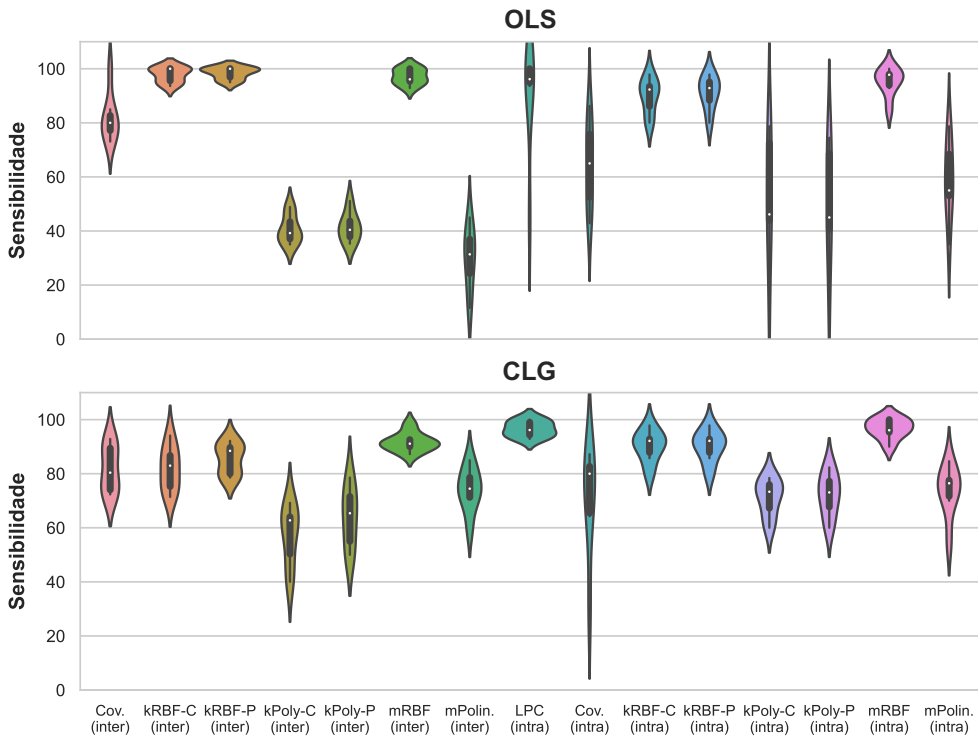
Ao analisar as sensibilidades obtidas pelos classificadores utilizados, é possível verificar através da Figura 44 que a utilização do CLG neste caso, causa um aumento na taxa de sensibilidade e uma menor dispersão para aqueles métodos em que se utilizou a matriz de kernel polinomial. Em contrapartida, ao utilizar a matriz de kernel RBF de modo intercanais o algoritmo de classificação OLS realiza previsões de sinais convulsivos melhor do que pelo CLG. No cenário intracanal, percebe-se que apesar de o método LPC ter obtido a maior média geométrica, este tem uma alta dispersão nas sensibilidades. Tal ocorrido é diferente nos métodos que utilizam a matriz de kernel, já que estes além de possuírem altos valores, tem uma baixa dispersão.

Figura 43 – **CHB01** Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.



Fonte: Elaborado pelo Autor.

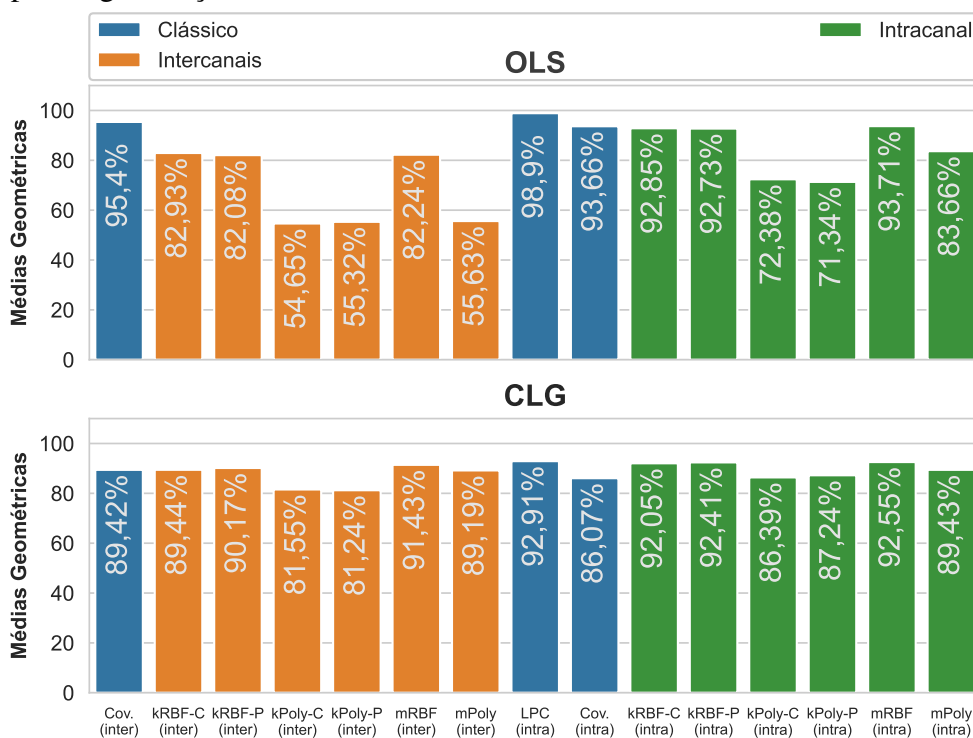
Figura 44 – **CHB01** Comparativo das sensibilidades entre OLS e CLG para segmentação 03.



Fonte: Elaborado pelo Autor.

Para a hipótese de segmentação 04, verificou-se ao analisar as médias geométricas, um aumento considerável para o método da covariância intercanais e o LPC, como pode ser visualizado através da Figura 45. Além disto, para a presente segmentação, foram observados melhores resultados comparados com as demais segmentações ao utilizar a matriz de kernel RBF e a abordagem intracanal. Contudo, a característica para o kernel polinomial pela abordagem intercanais, se mantém insatisfatória.

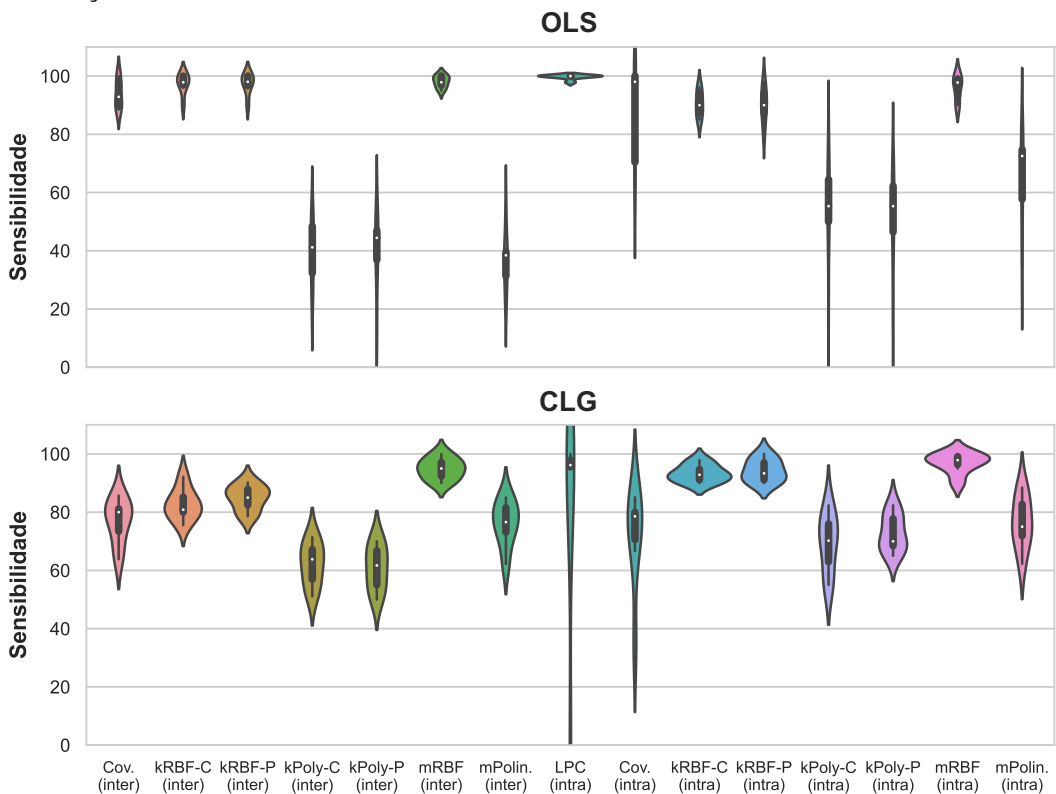
Figura 45 – **CHB01**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.



Fonte: Elaborado pelo Autor.

Ao analisar os gráficos exibidos através da Figura 46 pode-se dizer que a capacidade de prever corretamente os sinais epiléticos, é baixa para os métodos em que se utilizou a matriz de kernel polinomial. Além disto, percebe-se que os resultados obtidos na classificação via algoritmo dos mínimos quadrados e parametrização pelo método LPC concordam com a média geométrica exibida anteriormente. A análise realizada na parametrização que se utiliza da matriz de kernel RBF é similar ao que já foi discutido para as demais segmentações. Ou seja, a abordagem intracanal tem uma taxa de sensibilidade tão boa quanto aos métodos clássicos, e é superior quando se utiliza da esparsificação pela maximização da entropia de Renyi.

Figura 46 – **CHB01**: Comparativo das sensibilidades entre OLS e CLG para segmentação 04.



Fonte: Elaborado pelo Autor.

Para finalizar a análise do caso **CHB01**, pensou-se em destacar métodos de parametrização propostos que obtiveram os melhores resultados para cada uma das segmentações. Compôs-se então o Quadro presente no Apêndice A que exibe as taxas médias de acurácia, sensibilidade, especificidade e a média geométrica. Pelos resultados exibidos, pode-se reafirmar que os bons desempenhos (destacados em **negrito**) são obtidos ao utilizar a parametrização pela abordagem intracanal e matrizes de kernel RBF independente qual seja a esparsificação aplicada. Além disto, verifica-se que o kernel polinomial não teve resultados promissores, com exceção quando se faz a segmentação 04, que as médias geométricas obtidas são superiores aos métodos clássicos.

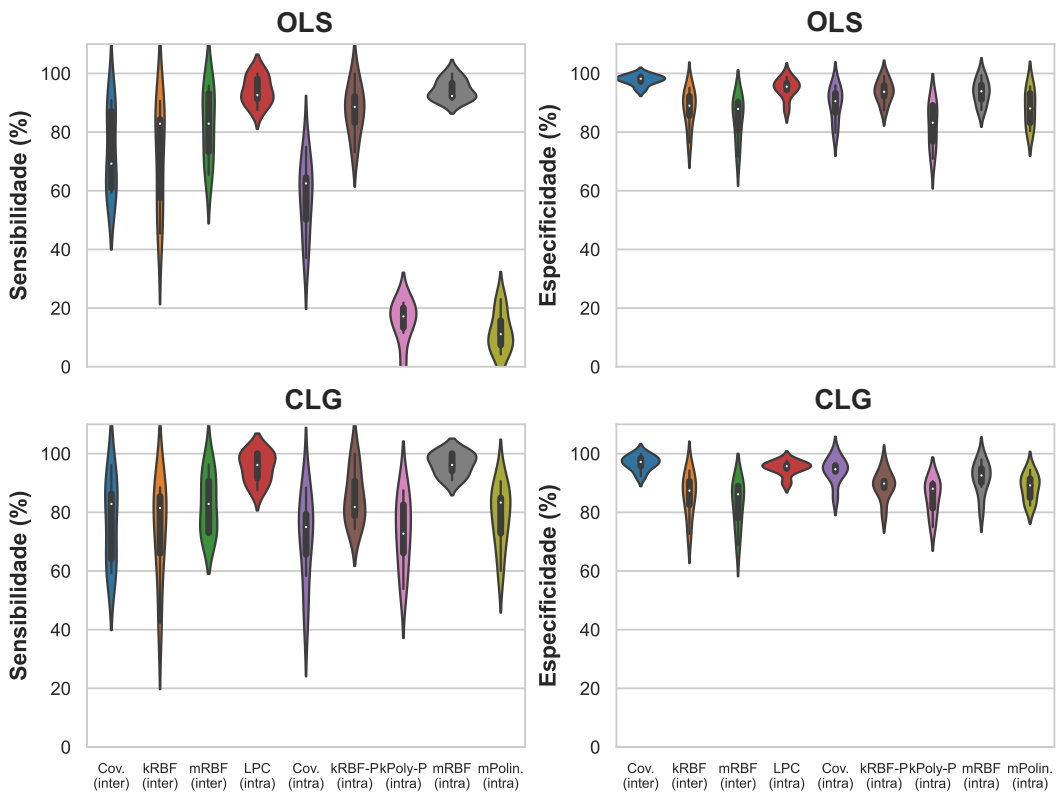
6.1.2 Resultados caso **CHB03**

Nesta subseção são exibidos os resultados obtidos ao analisar os arquivos do caso **CHB03**. Ressalta-se que a partir deste caso, alguns resultados são omitidos. Isto foi realizado, pois ao observar os resultados em que se fez o uso da abordagem intercanais em conjunto com a matriz de kernel polinomial, para qualquer uma das esparsificações, se obteve valores inferiores

aos demais métodos utilizados. Além destes mencionados, são omitidos também aqueles métodos em que se utilizou da esparsificação pelo k —Médias pela composição do dicionário de vetores relevantes ao selecionar os k vetores protótipos. Isto foi realizado, pois, para todas as análises os resultados são similares ou superiores quando se escolheu os k vetores mais próximos dos protótipos para compor o dicionário de vetores relevantes.

O caso **CHB03** dispõe de sete diferentes aquisições com uma hora de duração. Em média, há o registro de 57,42s com sinais epiléticos registrados. Ao aplicar os métodos nos arquivos deste caso para a hipótese de segmentação 01, pôde-se construir o gráfico da Figura 47. Nesta é possível verificar que além da informação de sensibilidade, exibem-se as especificidades obtidas ao executar cada método. Pode-se constatar ao observar estes resultados que para esta hipótese, o LPC possui um destaque com relação aos demais métodos. Porém, percebe-se que ao utilizar a esparsificação pelo critério de maximização da entropia de Renyi em conjunto com a matriz de kernel RBF (mRBF intra), os resultados obtidos são similares ao LPC.

Figura 47 – **CHB03** Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 01.

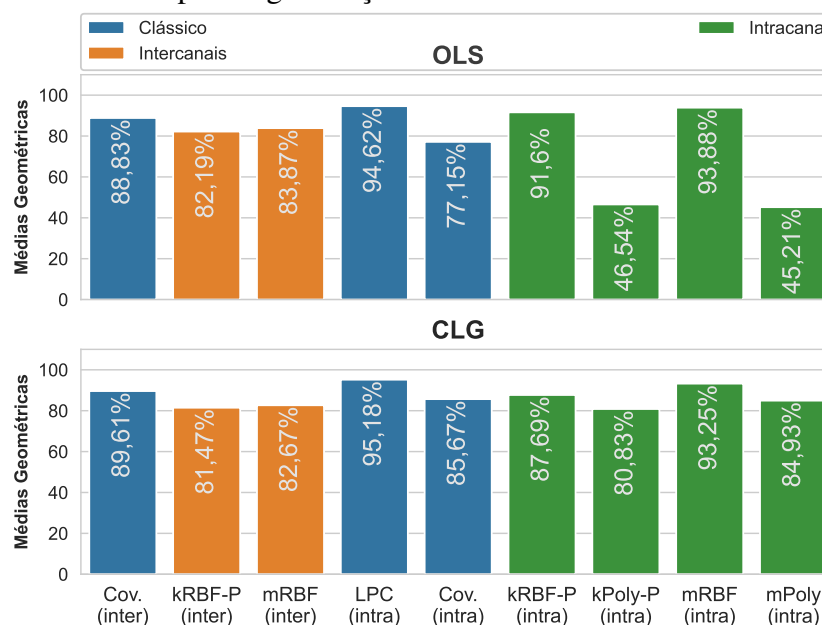


Fonte: Elaborado pelo Autor.

Analisando os resultados obtidos pelos métodos na abordagem intercanais, percebe-se que as sensibilidades são dispersas e inferiores aos melhores métodos. Porém, a especificidade da covariância intercanais aparenta possuir um desempenho melhor entre todos os métodos. Vale a pena lembrar que o problema em questão possui classes extremamente desbalanceadas e o ocorrido pode mascarar a taxa de acurácia. Ademais, as sensibilidades obtidas ao aplicar o kernel polinomial são perceptivelmente inferiores, mesmo quando utilizado o CLG.

Após a análise das sensibilidades e especificidades, compôs-se o gráfico ilustrado pela Figura 48, que exibe as médias geométricas obtidas para cada método. Pode-se perceber que as medidas exibidas são condizentes com a análise realizada anteriormente pela Figura 47, portanto, as discussões realizadas são similares. O método LPC possui a maior das médias geométricas, os métodos mRBF (intra) e kRBF-P (intra) possuem resultados promissores, porém inferiores ao LPC. Com respeito aos métodos que se utilizam da abordagem intercanais, destaca-se que a matriz de covariância tem resultado inferior aos métodos mRBF (intracanal) e kRBF-P (intracanal), porém, possui melhor resultado que os métodos mRBF (intercanais) e kRBF-P (intercanais), que obtiveram resultados não tão satisfatórios. Por fim, percebeu-se na aplicação do kernel polinomial, que os resultados obtidos são bem inferiores quando se utiliza o algoritmo OLS. Todavia, a aplicação do CLG nestes métodos, faz com que seu desempenho seja consideravelmente aumentado.

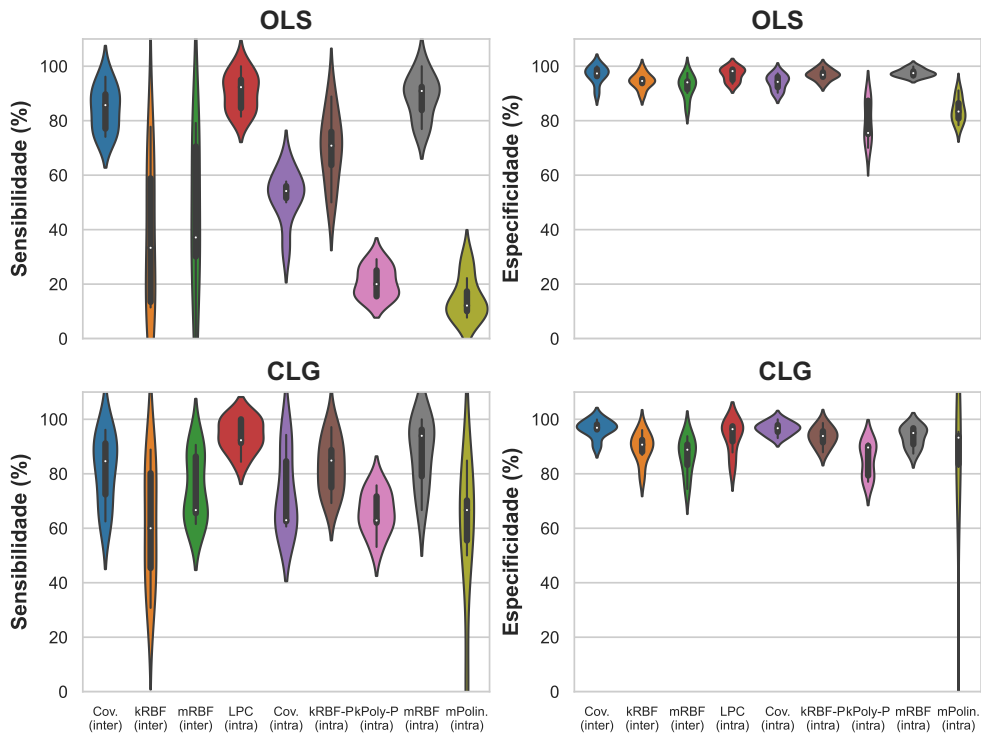
Figura 48 – **CHB03**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.



Fonte: Elaborado pelo Autor.

Tratando da segmentação pela hipótese 02, pode-se verificar através da Figura 49 que o método LPC possui o melhor dos desempenhos na tarefa de prever corretamente para as duas classes. Em sequência, os métodos mRBF (intracanal) e covariância (intercanais) apresentam resultados tão bons quanto o LPC apesar de inferior. Além disto, verifica-se que o método kRBF-P(intra) teve uma queda de valor comparando-se com a hipótese de segmentação 01, porém, o mesmo ainda obteve sensibilidades similares aos melhores métodos. Através desta Figura é possível dizer que os métodos que utilizam a matriz de kernel polinomial, caracterizam resultados bem inferiores aos demais quando aplicados na classificação via OLS. Este ocorrido é observado similarmente para o CLG, que apesar de um certo aumento de sensibilidade, ainda possui valores inferiores aos demais métodos.

Figura 49 – CHB03 Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 02.

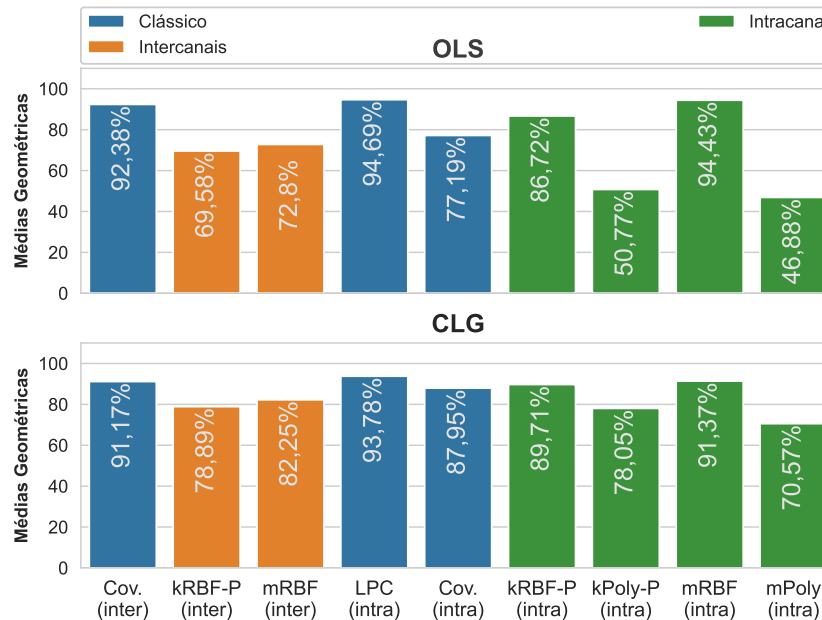


Fonte: Elaborado pelo Autor.

Pela análise das médias geométricas exibidas na Figura 50, pode-se perceber que os valores são condizentes às observações realizadas anteriormente. Destaca-se que utilizando a segmentação 02, os três métodos que possuem os maiores valores são o de covariância (intercanais), LPC e mRBF(intracanal). Pode-se visualizar que ao utilizar os demais métodos, os resultados foram bem inferiores ao comparar com os três melhores. Todavia, ressalta-se que o

CLG fez com que os piores resultados obtidos pelo OLS fossem aumentados consideravelmente, com destaque para aqueles que se utilizam da matriz de kernel polinomial.

Figura 50 – **CHB03** Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.

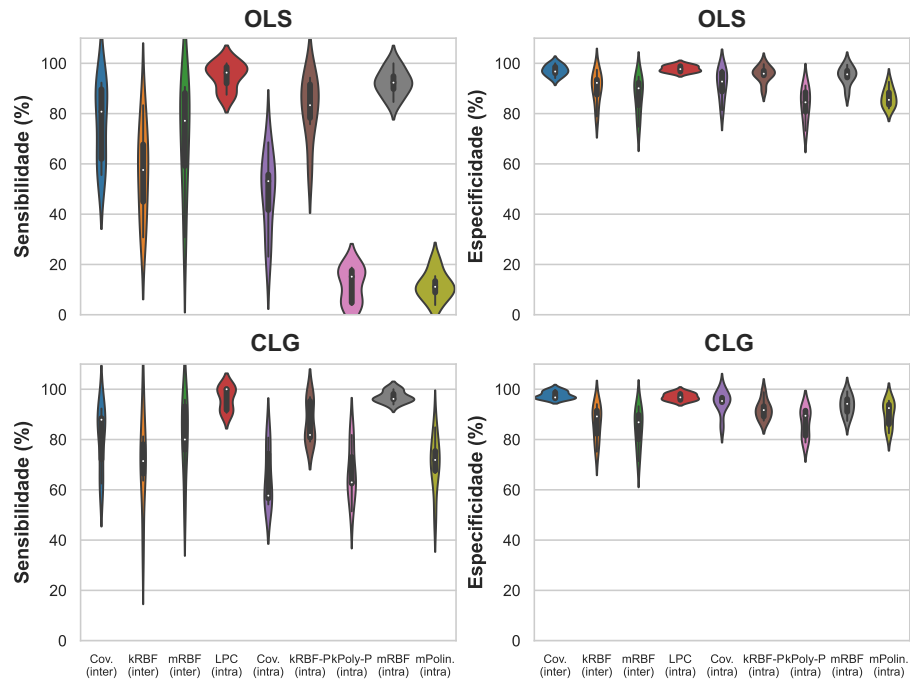


Fonte: Elaborado pelo Autor.

Para a hipótese de segmentação 03 que utiliza de amostras aleatórias do sinal não convulsivo, pôde-se perceber através da Figura 51 que a característica das sensibilidades e especificidades, são similares as segmentações anteriores. Ou seja, os métodos propostos de kRBF-P (intracanal) e mRBF(intracanal), possuem resultados inferiores aos métodos clássicos, porém, possuem valores próximos. Com respeito à sensibilidade obtida através dos dois algoritmos de classificação, destaca-se que o método proposto mRBF(intracanal) é tão bom quanto o LPC. Tal fato pode ser visualizado também na especificidade destes métodos e da covariância (intercanais). Ao realizar um comparativo entre os classificadores escolhidos, pode-se verificar que a dispersão das séries de resultados que utilizam da abordagem intracanal, é superior quando se aplica o método OLS. Logo, os resultados de sensibilidade nesta configuração para o CLG caracterizam uma melhoria, até mesmo para os métodos que se utilizam a matriz de kernel polinomial.

As médias geométricas calculadas nesta etapa, são exibidas através da Figura 52, o método proposto que utiliza a abordagem intracanal e esparsificação pelo critério da maximização da entropia de Renyi em conjunto a matriz de kernel RBF (mRBF intracanal), possui o valor tão bom quanto o método clássico LPC. Pode-se perceber ao comparar os resultados entre

Figura 51 – **CHB03**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 03.



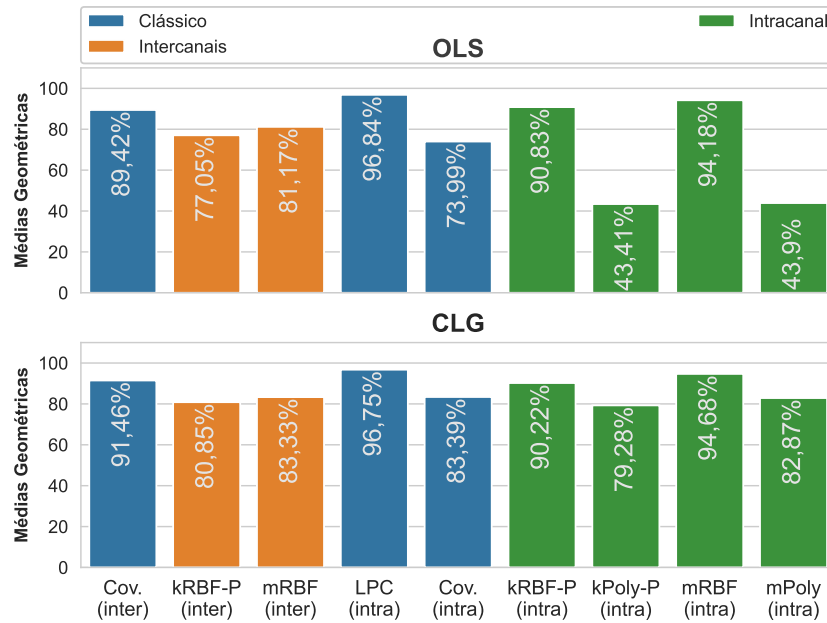
Fonte: Elaborado pelo Autor.

os classificadores que o CLG potencializa, em geral, as médias geométricas obtidas, porém, destacam-se os métodos que utilizam da matriz de kernel polinomial dos quais se observou uma maior elevação de resultados, apesar de ainda serem geralmente inferiores aos demais métodos.

Para concluir os resultados do caso **CHB03**, exibem-se as métricas de sensibilidade e especificidade através da Figura 53 e médias geométricas através da Figura 54. Pode-se concluir com as análises realizadas que apesar de se ter segmentado o sinal por diferentes maneiras, os resultados obtidos neste caso foram similares. Para todas as segmentações o método clássico LPC teve um maior desempenho, com valores acima de 90% de média geométrica. Em sequência se encontra o método proposto mRBF(intracanal) que apesar de ter médias geométricas inferiores ao LPC, também possui valores acima de 90%.

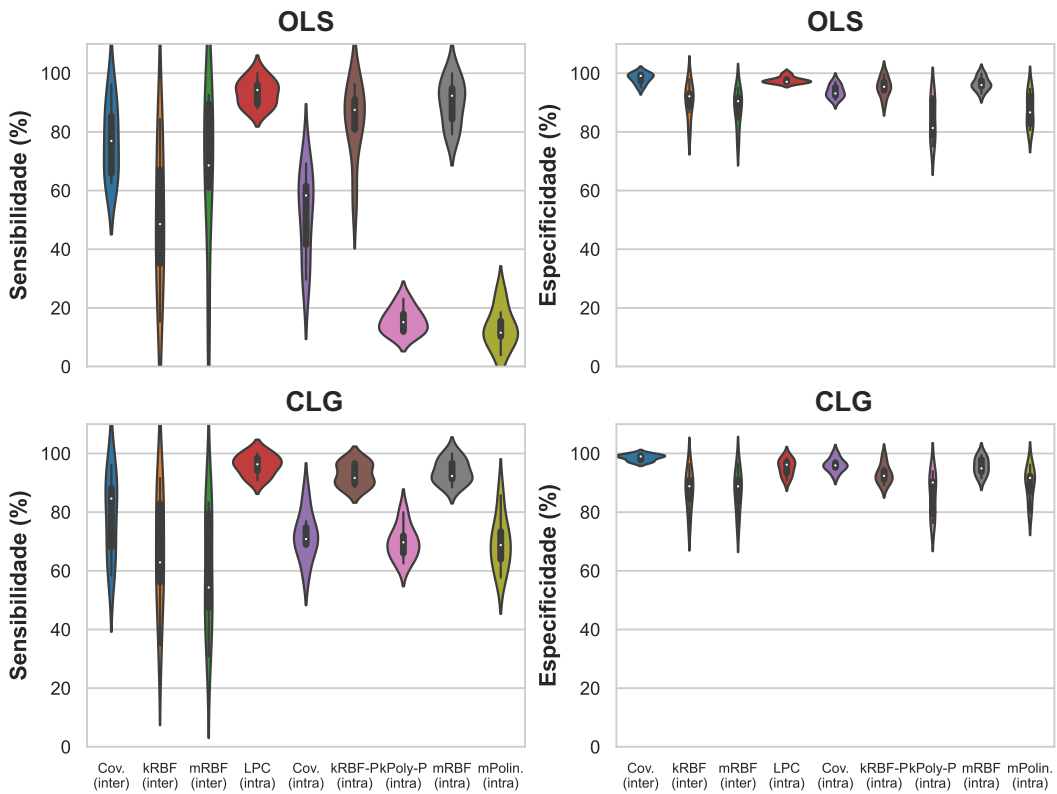
O método de covariância (intercanais) se apresenta como terceiro melhor método para o caso **CHB03** e dependendo da segmentação, possui médias geométricas acima de 90% também. Destaca-se ainda o método proposto kRBF-P (intracanal), possui valores próximos ao método de covariância, mesmo que inferiores em certas segmentações. Uma análise importante com relação aos resultados obtidos no presente caso, é de que o método da covariância teve resultados de sensibilidade baixos comparando-se aos dois métodos propostos que tiveram maiores médias geométricas. Logo, o método de covariância (inter) obteve, para alguns casos,

Figura 52 – **CHB03** Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.



Fonte: Elaborado pelo Autor.

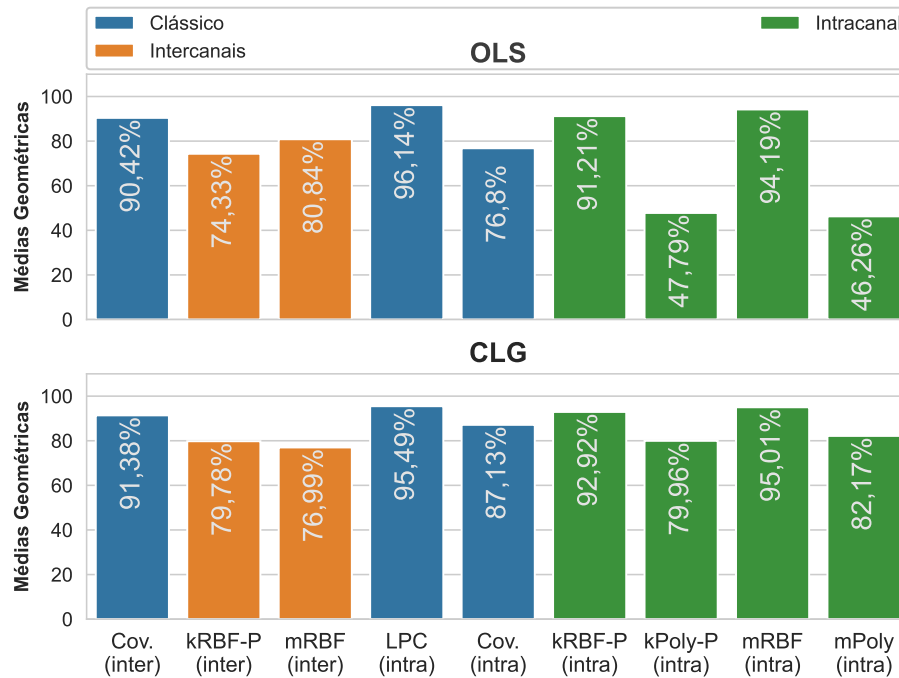
Figura 53 – **CHB03**: Comparativo das sensibilidades entre OLS e CLG para segmentação 04.



Fonte: Elaborado pelo Autor.

a média geométrica acima de 90%, devido ao seu bom desempenho na taxa de especificidade. Por fim, disponibiliza-se no Anexo B, o Quadro 6, que traz informações dos valores médios de acurácia, sensibilidade e especificidade em cada uma das segmentações.

Figura 54 – **CHB03**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.



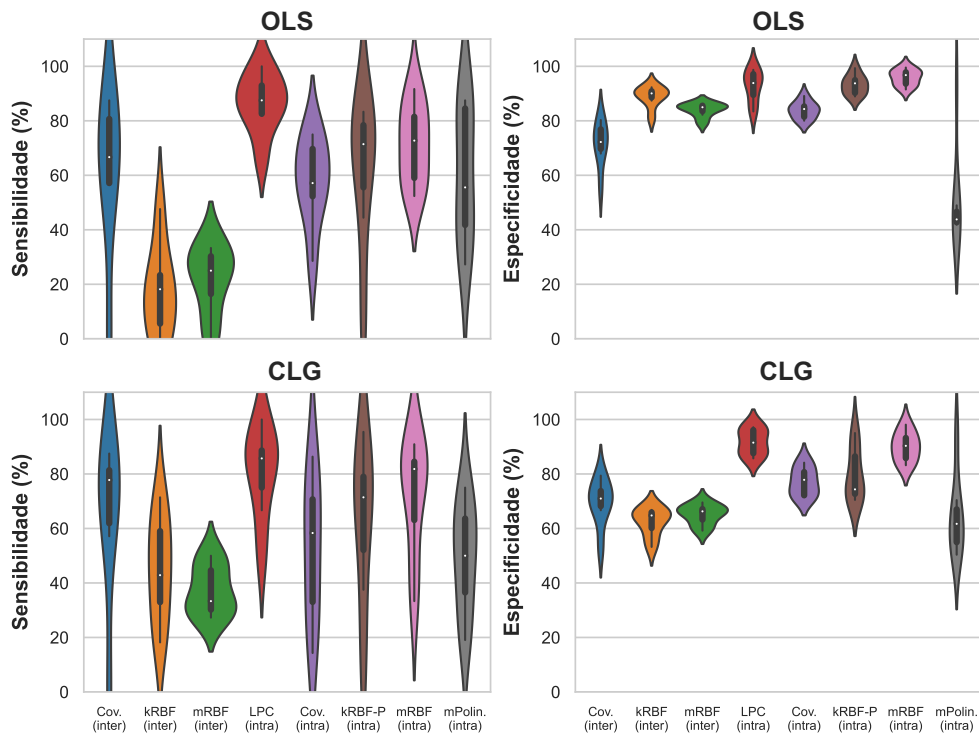
Fonte: Elaborado pelo Autor.

6.1.3 Resultados do caso CHB14

As aquisições em que se foi identificado a presença de atividades convulsivas no paciente do caso **CHB14**, foram armazenadas em sete arquivos com duração de uma hora cada. Destas 7 horas, apenas 147s são caracterizados como sinais epiléticos e em média, há $21 \pm 9,36s$ por arquivo. Inicialmente fez-se a aplicação dos métodos pela hipótese de segmentação 01 e foi possível construir o gráfico da Figura 55, que exhibe as taxas de sensibilidade e especificidade obtidas. Nesta etapa, foi possível verificar que geralmente, os resultados obtidos são inferiores aos casos anteriores, por conta dos gráficos possuírem alta dispersão e uma menor mediana.

Ao analisar as sensibilidades obtidas, verifica-se que o método LPC possui certa superioridade com relação aos demais. Este ocorrido também pode ser visualizado ao observar os resultados exibidos para especificidade. Os métodos que se destacam em sequência são os propostos pelo presente trabalho (kRBF-P e mRBF) que utilizam da abordagem intracanal. Estes

Figura 55 – CHB14: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 01.



Fonte: Elaborado pelo Autor.

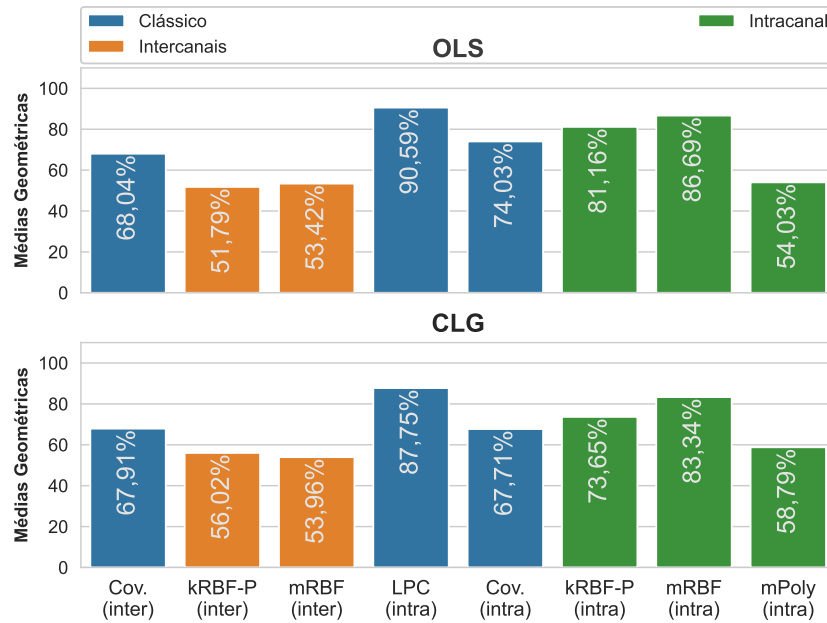
possuem resultados tão bons ou até melhores que o LPC com respeito a especificidade, porém, perdem a capacidade de realizar predições da condição positiva.

Observando as médias geométricas obtidas nesta segmentação, confirmam-se as análises realizadas anteriormente. Portanto, como se pode visualizar através da Figura 56 os métodos que se destacam são o LPC, o kRBF-P e o mRBF, ambos pela abordagem intracanal. Os demais métodos tiveram valores de médias geométricas inferiores a 80%. Percebe-se ainda que para exemplos como os que utilizam a matriz de kernel polinomial e os que fazem uso da abordagem intercanais, tendem a resultados não promissores.

Realizando a segmentação pela hipótese 02, foi constatado que o desempenho dos métodos clássicos tiveram um decaimento, com destaque para o LPC que como pode ser visualizado através da Figura 57, obteve taxas de sensibilidade dispersas e especificidade baixas. Nesta Figura ainda é possível verificar que os métodos propostos que tiveram maior desempenho na segmentação anterior, tiveram taxas similares para a presente segmentação. Apesar de se ter constatado que suas sensibilidades são inferiores às obtidas pelo LPC, as especificidades no que lhe concernem, apresentam as maiores taxas entre todos os métodos analisados.

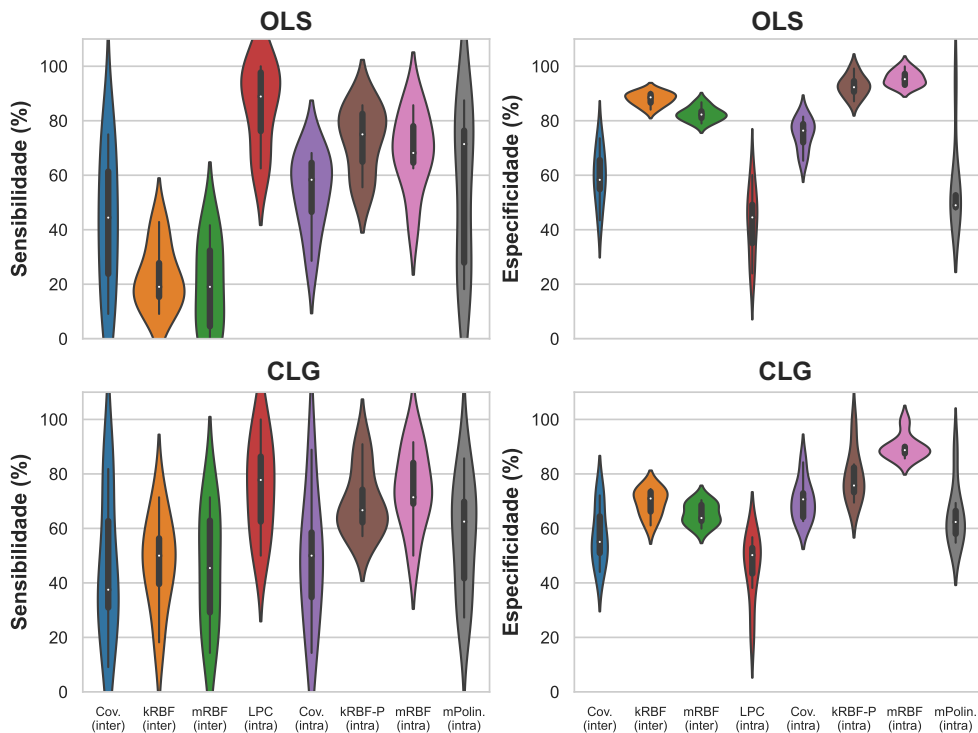
Com respeito aos resultados referentes as médias geométricas, pode-se perceber ao

Figura 56 – **CHB14** Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.



Fonte: Elaborado pelo Autor.

Figura 57 – **CHB14**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 02.

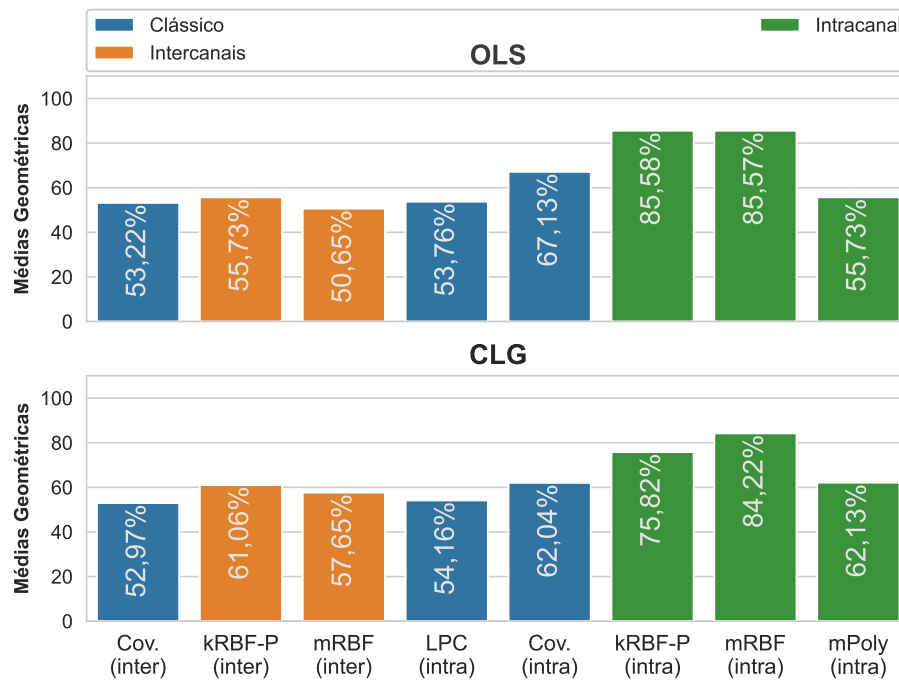


Fonte: Elaborado pelo Autor.

visualizar os dados na Figura 58 que os métodos utilizados têm valores abaixo de 70%, com exceção dos propostos kRBF-P e mRBF aplicados pela abordagem intracanal, que se mantiveram

num patamar acima de 85% quando se aplicou o classificador OLS. Destaca-se também que a utilização do CLG para esta configuração, em geral, não proporcionou melhorias suficientes para ultrapassar dos melhores exemplos do OLS.

Figura 58 – **CHB14**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.

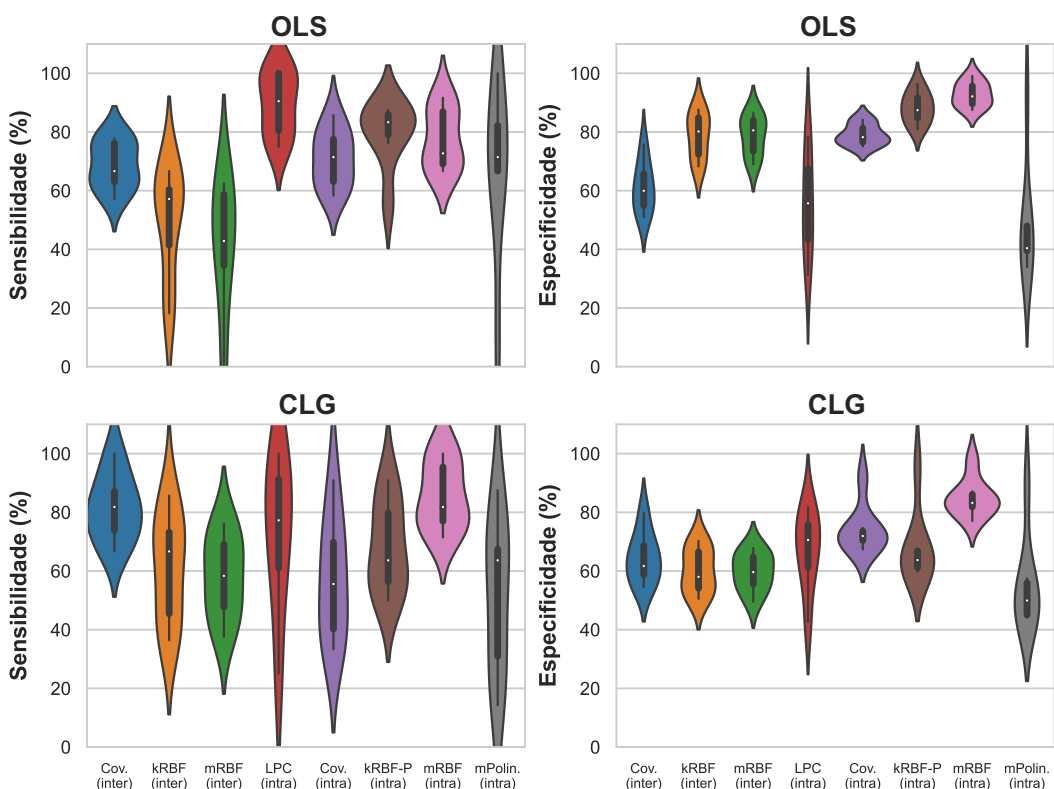


Fonte: Elaborado pelo Autor.

Para a hipótese de segmentação 03, ao realizar a rodada de testes com os métodos utilizados, pôde-se compor o gráfico exibido através da Figura 59. Tratando das sensibilidades exibidas nesta Figura, é possível dizer que os métodos clássicos que fazem o uso da matriz de covariância, não são tão bons quanto aos resultados obtidos pelo método clássico LPC. Este no que lhe concerne apresenta a gama de melhores sensibilidades (para o OLS) com relação aos demais métodos. Destaca-se também que com relação aos métodos propostos, os que fazem o uso da abordagem intracanal tem desempenhos melhores do que pela intercanais. Faz-se necessário comentar que o uso do classificador Linear Gaussiano, eleva as sensibilidades destas últimas abordagens, com destaque para o método mRBF (intracanal), que possui melhores taxas pelo CLG.

Com relação às especificidades, destacam-se que os métodos propostos pela abordagem intracanal e que se utilizam da matriz de kernel RBF, possuem os valores mais altos e com menor dispersão considerando o classificador OLS. Porém, ao analisar os dados equivalentes

Figura 59 – **CHB14**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 03.



Fonte: Elaborado pelo Autor.

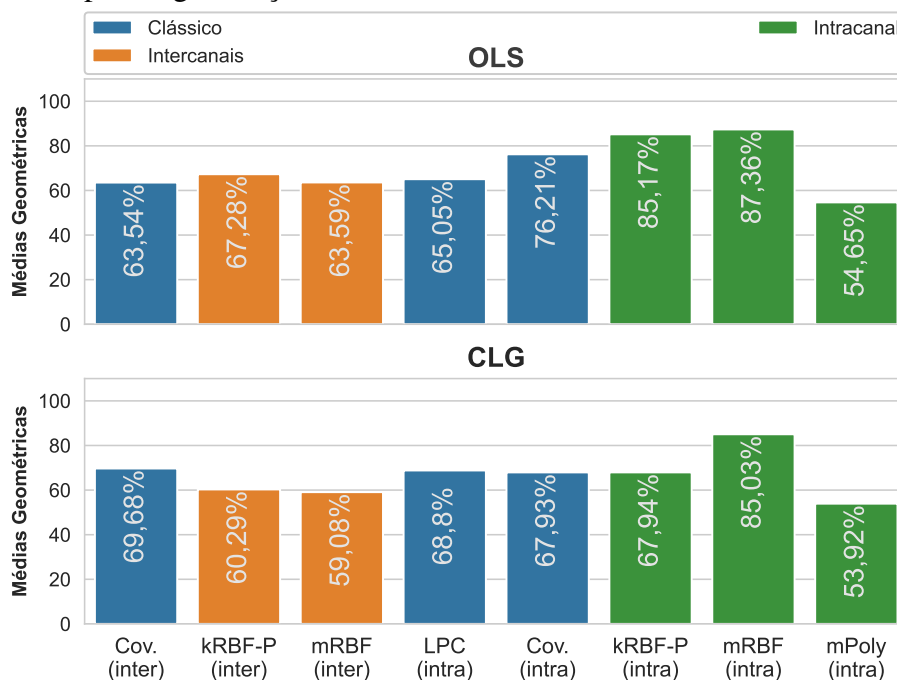
obtidos pelo CLG, percebe-se um decaimento em geral nas taxas obtidas.

A análise que pode ser feita ao visualizar as médias geométricas na Figura 60, é que os métodos propostos kRBF-P e mRBF ambos pela abordagem intracanal, têm as maiores taxas. Em sequência encontra-se o método clássico LPC, que apesar de realizar previsões mais corretas para a condição positiva, não faz o mesmo para a condição negativa. Confirma-se também através da Figura que a aplicação do CLG nesta rodada proporcionou um desempenho inferior comparando-se ao classificador OLS, com exceção de quando se utilizou o método LPC.

Ao realizar a segmentação pela hipótese 04, os resultados exibidos pela Figura 61 são similares a primeira segmentação. O método LPC possui sensibilidades superiores aos demais métodos. Em sequência encontram-se os métodos propostos kRBF-P e mRBF, ambos pela abordagem intracanal. Porém, ao analisar as especificidades obtidas pelo classificador OLS, verifica-se que a dispersão das taxas no método mRBF (intra) é menor que no método LPC.

Comparando os resultados obtidos através dos dois algoritmos de classificação, ressalta-se que o uso CLG proporcionou um aumento das sensibilidades nos métodos em que o OLS obteve menores resultados. Porém, nos demais métodos as sensibilidades obtidas pelo

Figura 60 – **CHB14**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.



Fonte: Elaborado pelo Autor.

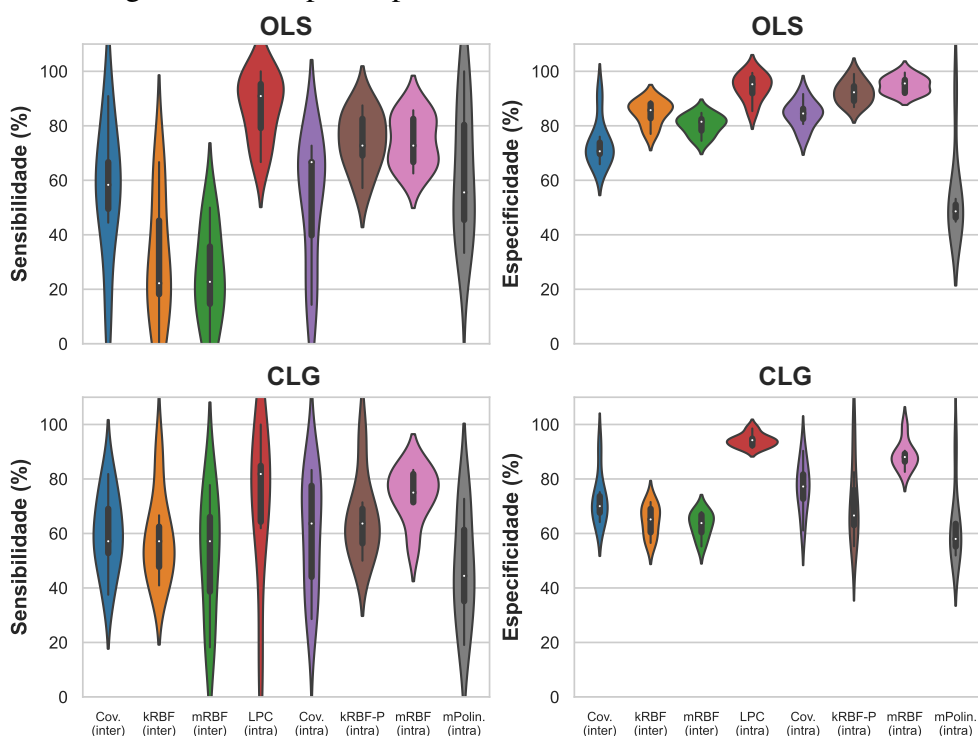
CLG são inferiores. Com relação à especificidade, essa análise é diferente, pois aparentemente há apenas uma leve melhoria ao utilizar o na classificação pelo CLG com o método LPC.

Com o cálculo das médias geométricas realizadas, compôs-se o gráfico ilustrado pela Figura 62 e foi possível confirmar o discutido sobre esta segmentação. O método clássico LPC possui o maior dos resultados com valor de 91,56% em sequência se encontram os métodos propostos kRBF-P (intra) e mRBF (intra) com respectivamente 85,79% e 87,42%. Os demais métodos são considerados não satisfatórios pelo fato dos valores observados serem abaixo de 75%. Destaca-se que para este caso **CHB14**, os resultados obtidos ao parametrizar o sinal utilizando a matriz de kernel polinomial são insatisfatórios e no geral tem menores valores entre todos os métodos. Por fim, disponibiliza-se no Anexo C, o Quadro 7, que traz informações dos valores para as demais segmentações.

6.1.4 Resultados caso CHB18

Após a exibição dos resultados para os casos anteriores, pode-se perceber que a utilização da matriz de kernel polinomial para o presente conjunto de dados, foi insatisfatória. Tal fato também pode ser verificado nos resultados obtidos pelos casos que estão disponíveis no Apêndice F. Logo, para os dois últimos casos, também são omitidos os resultados do método

Figura 61 – **CHB14**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 04.



Fonte: Elaborado pelo Autor.

Figura 62 – **CHB14**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.

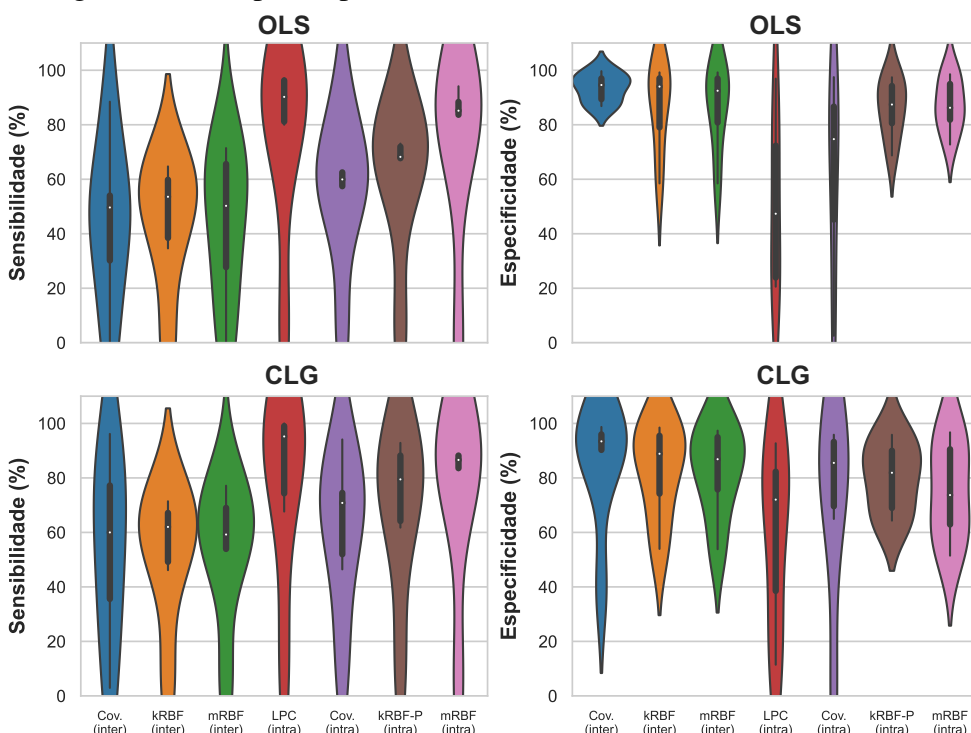


Fonte: Elaborado pelo Autor.

proposto através da esparsificação pelo critério de entropia de Reniy em conjunto com a matriz de kernel polinomial.

O caso **CHB18** foi escolhido de modo a representar os outros casos em que os resultados em geral são baixos. Provavelmente a razão deste ocorrido, se dá pela presença de artefatos nas aquisições dos sinais de EEG. O presente caso possui seis arquivos com períodos convulsivos identificados, e possui uma média de 52,83s destes sinais por arquivo. Em uma análise inicial, pode-se perceber através da Figura 63 que as sensibilidades obtidas são geralmente menores com relação aos outros casos já descritos. Todavia, o método LPC apesar de ser bem disperso, possui taxas mais próximas de 100%. A dispersão das sensibilidades também é observada para todas as configurações. Realizando a análise das especificidades, verifica-se que a maioria dos métodos tem um desempenho satisfatório, porém, os métodos LPC e covariância (intra) são bem inferiores. Ao levantar o comparativo entre os métodos de classificação utilizados, percebe-se que, em geral, ambos possuem desempenhos similares. Faz-se um destaque apenas para os métodos LPC e covariância (intercanais) que obtiveram uma melhoria nas duas taxas quando foram classificados pelo CLG.

Figura 63 – **CHB18** Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 01.



Fonte: Elaborado pelo Autor.

Condensando as métricas obtidas, nas médias geométricas, pôde-se compor o gráfico na Figura 64, na qual é possível constatar que os dois métodos propostos pela abordagem intracanal têm taxas superiores aos demais métodos. Em sequência, encontra-se o método clássico de covariância (inter) com valor de 72,79%. Ressalta-se que são poucas as diferenças entre a utilização do OLS e do CLG, porém, é possível destacar que o método que obteve a maior das médias geométricas, foi observado ao utilizar o classificador OLS.

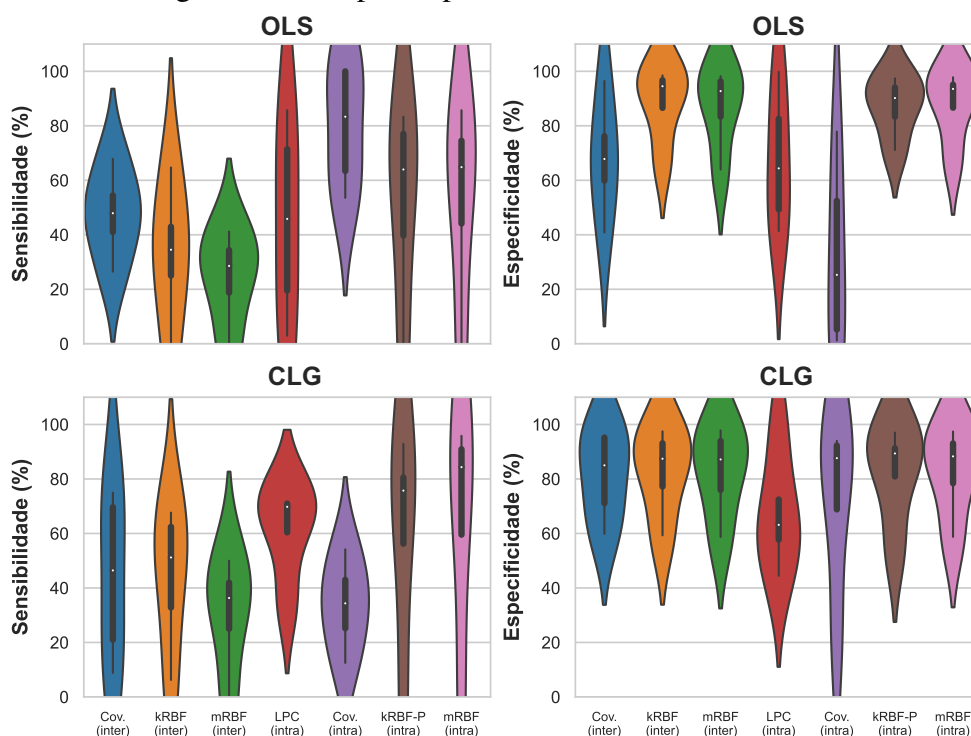
Figura 64 – **CHB18** Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.



Fonte: Elaborado pelo Autor.

Analisando os resultados obtidos pela aplicação da hipótese 02 de segmentação, percebeu-se que todas as taxas de sensibilidades e especificidades tem desempenho inferior pelas análises realizadas na segmentação 01. Porém, ainda assim, como pode ser visualizado através da Figura 65 as especificidades obtidas pelos métodos propostos no presente trabalho, são superiores que as obtidas pelos métodos clássicos. Isto pode ser visualizado também na sensibilidade obtida ao utilizar o CLG.

Figura 65 – **CHB18**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 02.

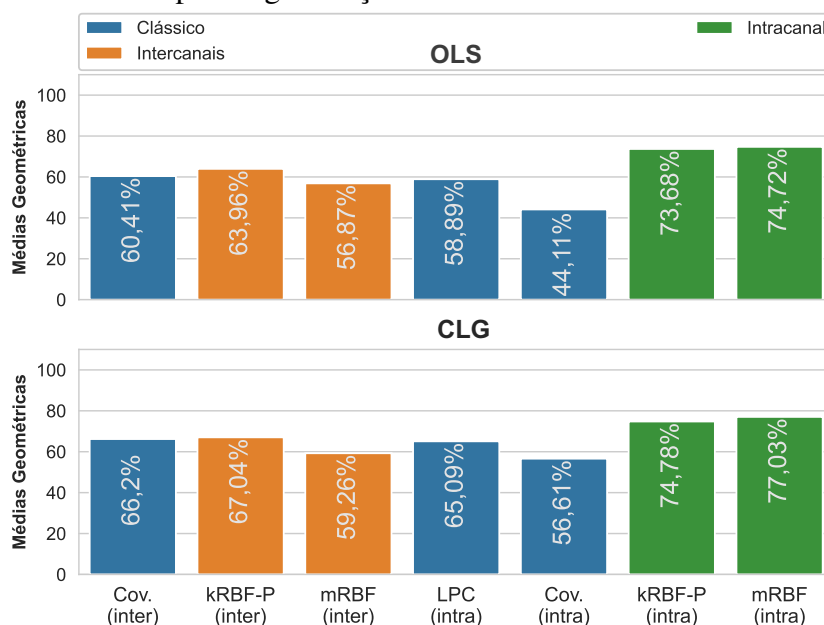


Fonte: Elaborado pelo Autor.

Realizando a análise conjunta através das médias geométricas, percebe-se através da Figura 66 que o método CLG nesta segmentação tem desempenho similar ao da segmentação 01. Isto se difere ao analisar estes resultados pelo classificador OLS, que tem resultados inferiores. Independente do classificador, pode-se afirmar que os métodos propostos no presente trabalho, que utilizam a abordagem intracanal, possuem taxas de médias geométricas mais elevadas com relação aos demais métodos de parametrização utilizados.

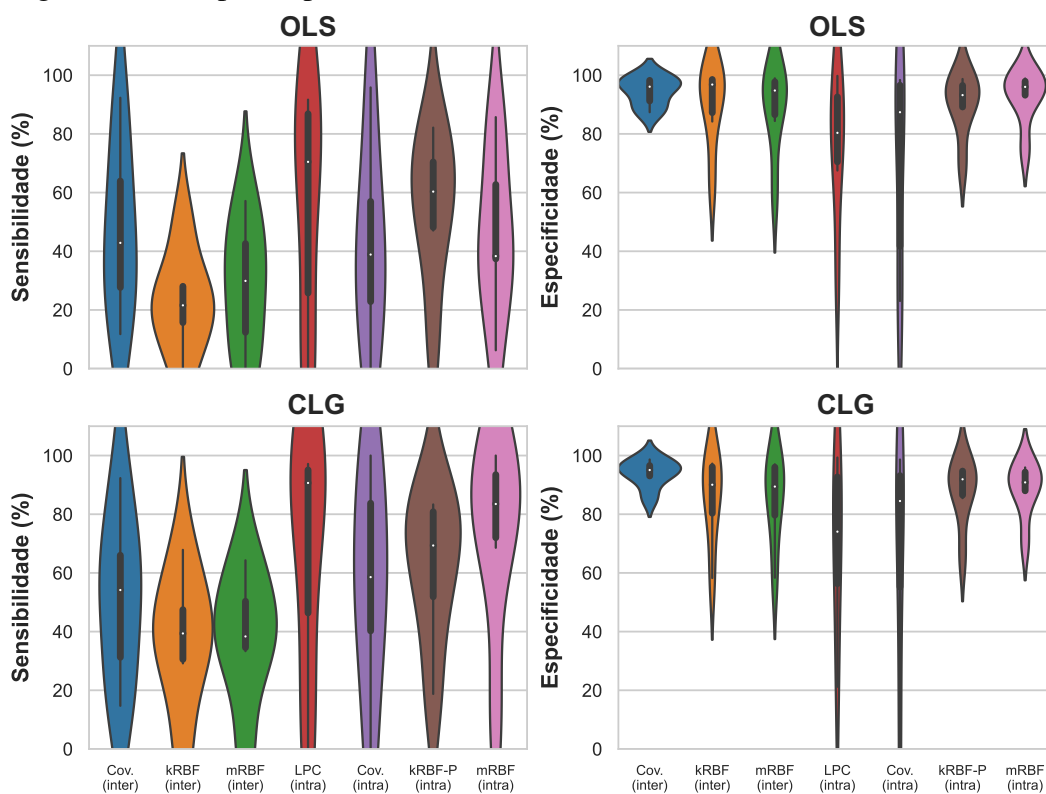
Como dito anteriormente, a segmentação pela hipótese 03 faz o uso de amostras aleatórias de sinais não convulsivos presentes nos arquivos de cada caso. Percebe-se ao analisar as sensibilidades e especificidades obtidas na etapa de classificação, exibidas pela Figura 67, que os desempenhos são em geral, similares aos obtidos pela segmentação 01. Quando se utiliza o algoritmo OLS, percebe-se que o método proposto kRBF-P (intra) tem desempenho superior ao realizar uma análise simultânea entre as sensibilidades e especificidades obtidas. Porém, pelo uso do CLG percebeu-se que o método proposto mRBF (intra) é que tem o maior dos desempenhos.

Figura 66 – **CHB18**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 02.



Fonte: Elaborado pelo Autor.

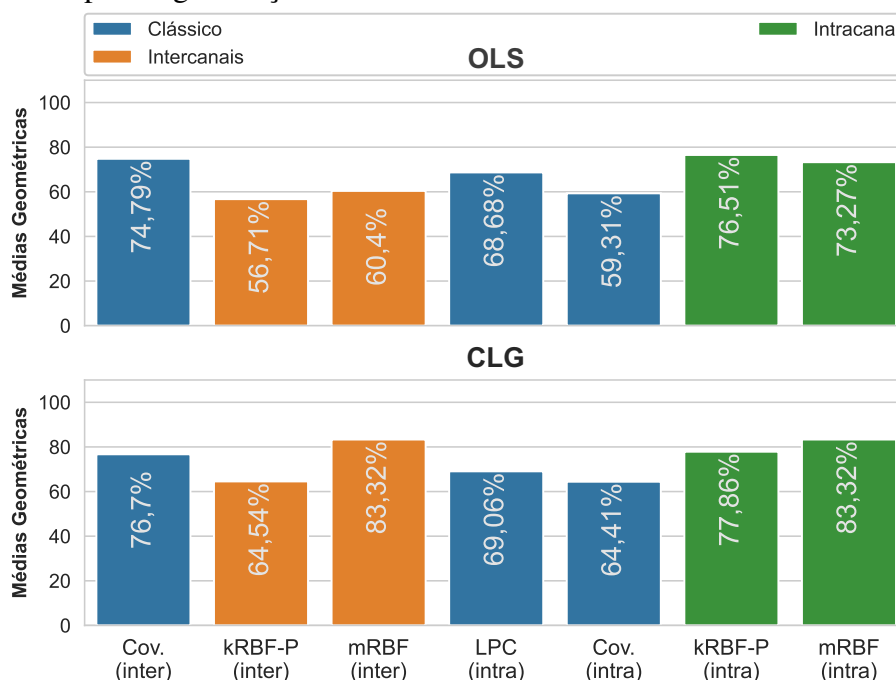
Figura 67 – **CHB18**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 03.



Fonte: Elaborado pelo Autor.

Essas informações de desempenhos são confirmadas pela análise de médias geométricas. Como pode ser visualizado através da Figura 68, os métodos propostos pela abordagem intracanal tem taxas superiores aos demais métodos. Além disto, pode-se destacar que ao utilizar o CLG, a parametrização via mRBF(intra) teve resultados promissores, tendo em vista que se obteve uma taxa acima de 80%.

Figura 68 – **CHB18**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.

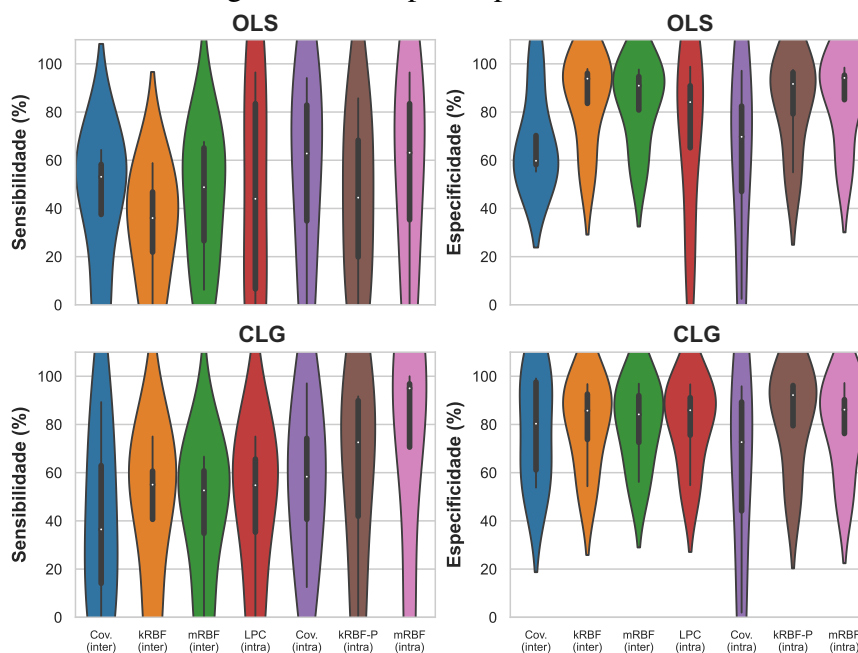


Fonte: Elaborado pelo Autor.

Aplicando-se a segmentação através da hipótese 04, percebe-se através da Figura 69 que as sensibilidades obtidas em todas as observações, são no geral, dispersas. Porém, destaca-se a parametrização mRBF(intra) que tem desempenho melhor que os demais métodos analisando a classificação pelo algoritmo do CLG. Tratando das especificidades, pode-se perceber que há uma menor dispersão e também uma elevação das medianas no CLG do que no OLS. Percebe-se pelo comportamento dos *violin plots*, que também o método mRBF(intra) tem superioridade aos demais.

Essa observação realizada, pode ser confirmada através da Figura 70 que exhibe os gráficos de médias geométricas para todos os métodos de parametrização. Portanto, verifica-se que o desempenho que é superior aos demais é observado no método que se utiliza da abordagem intracanal e realiza a esparsificação através do critério de maximização da entropia de Renyi

Figura 69 – **CHB18**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 04.



Fonte: Elaborado pelo Autor.

em conjunto da matriz de kernel RBF. Na presente segmentação, pode-se verificar que o uso do CLG proporcionou no geral um aumento nos valores das taxas obtidas. Por fim, pode-se verificar no Anexo D, demais resultados obtidos pelas segmentações através do Quadro 8.

Figura 70 – **CHB18**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.

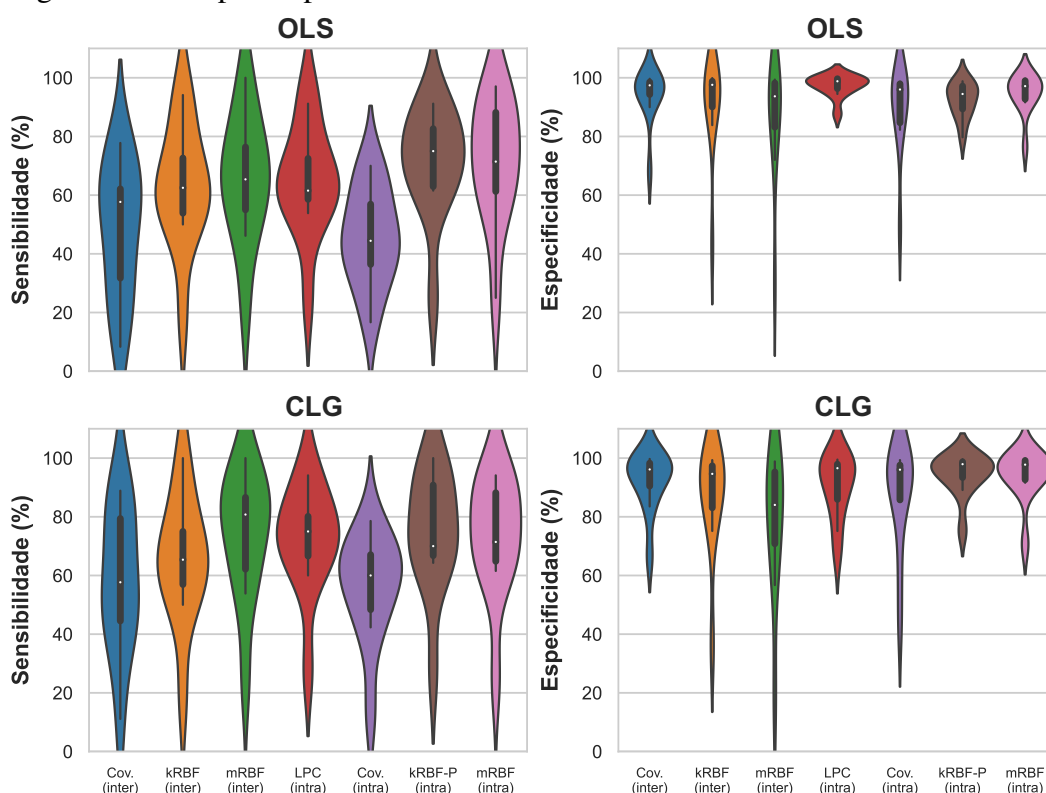


Fonte: Elaborado pelo Autor.

6.1.5 Resultados caso CHB24

Para compor o último resultado do primeiro conjunto de dados, fez-se a escolha do caso que represente o conjunto de casos, em que os desempenhos dos métodos propostos são sempre superiores aos clássicos. Portanto, escolheu-se o caso **CHB24** que possui 11 arquivos de aquisição do sinal e possui em média $33,57 \pm 19,18s$ de sinais epiléticos por arquivo. Inicialmente, foram realizadas as análises pela hipótese de segmentação 01 e pôde-se construir o gráfico exibido pela Figura 71. Nesta é possível verificar que as sensibilidades obtidas ao utilizar o método kRBF-P (intra) pela classificação no OLS, tem uma característica superior as demais parametrizações. Ao analisar as especificidades obtidas, constata-se que o método clássico LPC e o proposto mRBF (intra) possuem os valores de maior magnitude e uma dispersão aceitável.

Figura 71 – **CHB24**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 01.

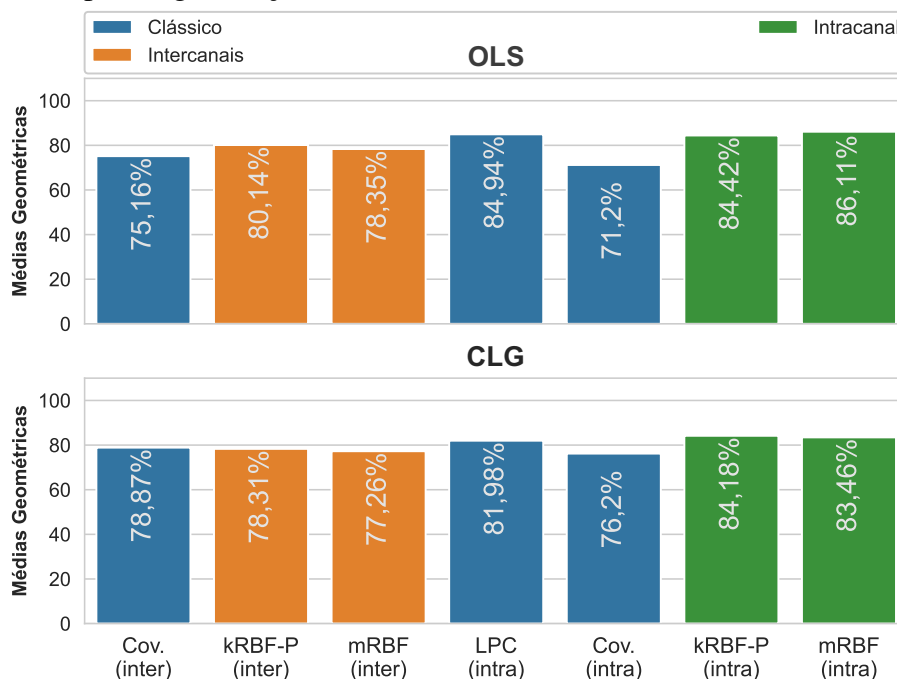


Fonte: Elaborado pelo Autor.

Como as características dos *violin plots* nos melhores métodos identificados são parecidas, fez-se o uso do cálculo de média geométrica entre acurácia, sensibilidade e especificidade. Desta maneira, pode-se perceber ao visualizar a Figura 72 que para a classificação via OLS, obtém-se como melhor método o mRBF (intra) com valor de 86,11%. Utilizando o CLG como

classificador, percebe-se que o método de maior média geométrica é o kRBF-P (intra). Portanto, nesta segmentação os melhores métodos são os propostos pela presente dissertação.

Figura 72 – **CHB24**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 01.



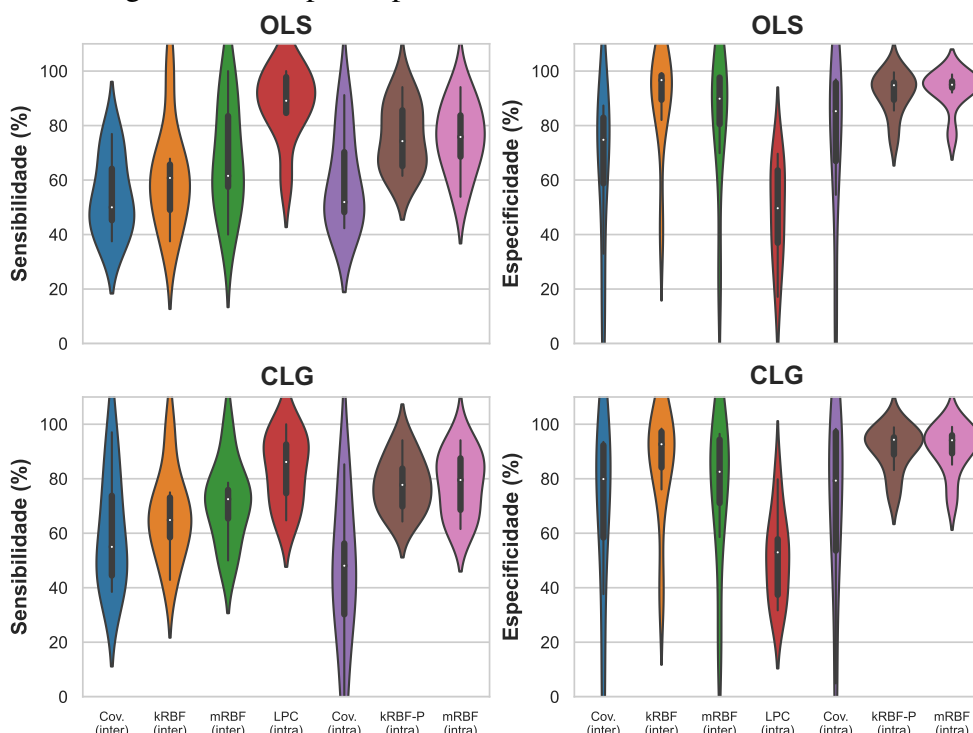
Fonte: Elaborado pelo Autor.

Os resultados obtidos pela hipótese 02 de segmentação do sinal, podem ser visualizados através da Figura 73. Pode-se constatar ao comparar as sensibilidades obtidas pelos métodos de classificação, que em ambas a parametrização via LPC têm uma gama de resultados superior aos demais métodos analisados, porém, quando essa análise é realizada na especificidade, o LPC possui resultados inferiores aos demais. Desta maneira, ao eliminar o método LPC, é possível visualizar que os dois métodos propostos pelo presente trabalho, que utilizam a abordagem intracanal, tem desempenhos superiores tanto na sensibilidade quanto na especificidade.

Tal ocorrido é verificado ao compor o gráfico da Figura 74. Pode-se observar que as médias geométricas superiores, são referentes aos métodos kRBF-P (intra) e mRBF (intra), com valores respectivamente de 86,03% e 86,82%. Constata-se também que os métodos clássicos nessa segmentação tiveram desempenho inferior aos métodos propostos. Ainda assim, os métodos que utilizam a abordagem intracanal, possuem uma magnitude de médias geométricas maiores do que os que se utilizam da abordagem intercanais.

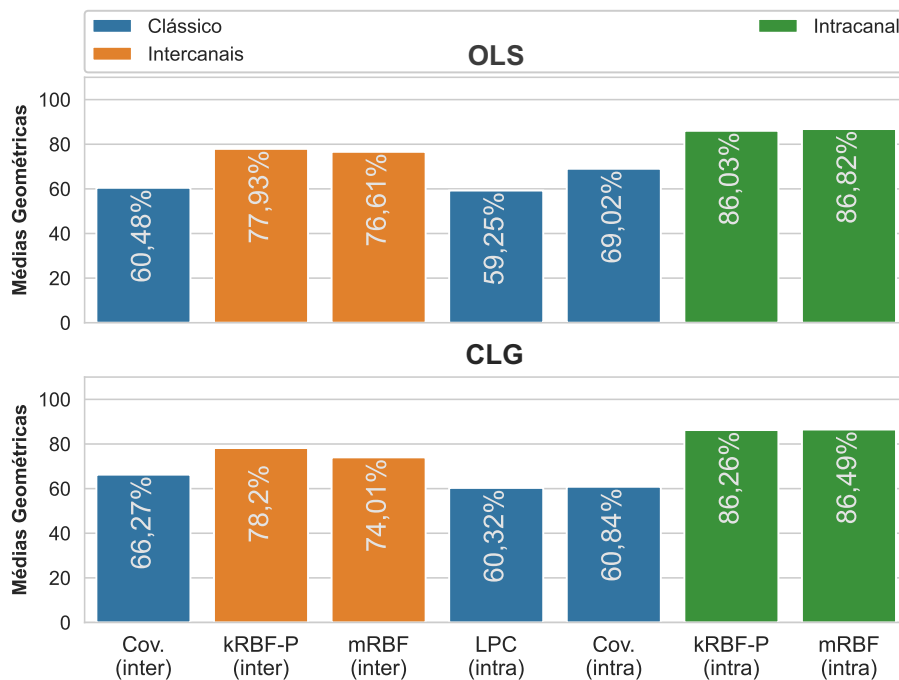
As sensibilidades e especificidades obtidas pelos algoritmos de classificação ao

Figura 73 – **CHB24**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese O2.



Fonte: Elaborado pelo Autor.

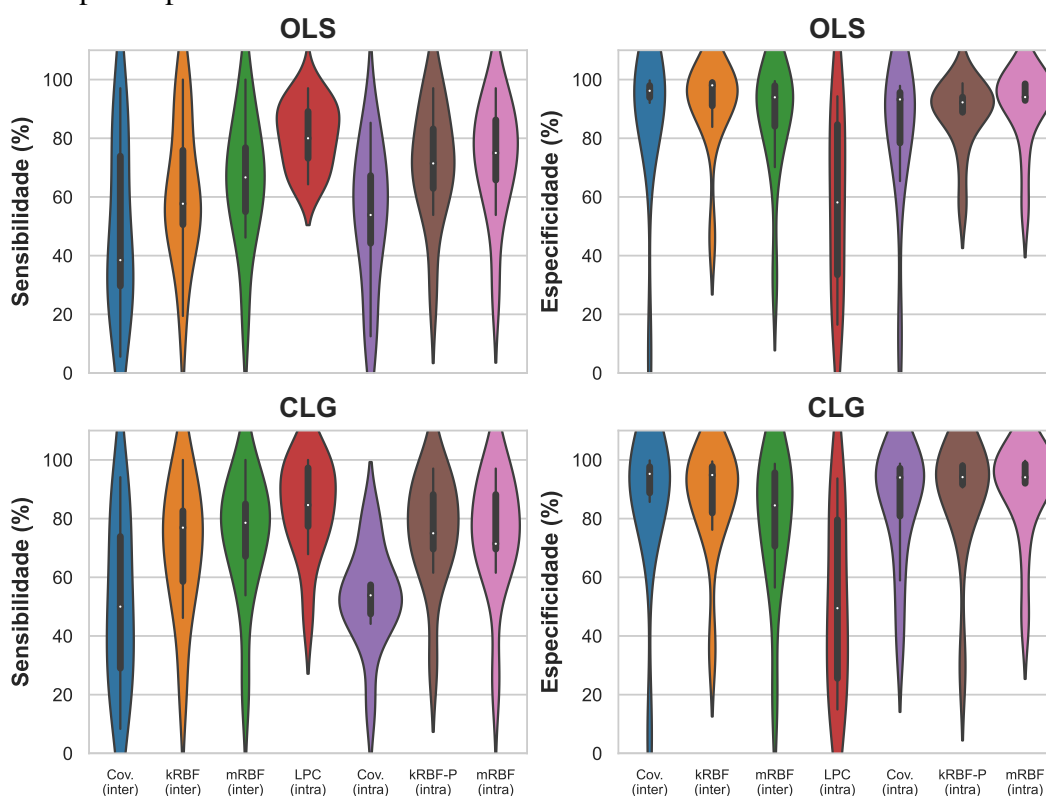
Figura 74 – **CHB24** Comparativo das Médias Geométricas entre OLS e CLG para segmentação O2.



Fonte: Elaborado pelo Autor.

utilizar a hipótese 03 de segmentação, podem ser visualizadas através da Figura 75. Verifica-se nesta, que o método clássico LPC se comportou similarmente ao descrito pela hipótese 02 de segmentação. Ou seja, fez boas previsões para a classe positiva e ruins para a classe negativa. Analisando os demais métodos clássicos pode-se identificar uma alta dispersão nas sensibilidades. Em contrapartida, os métodos propostos pelo presente trabalho que fazem o uso da abordagem intracanal, possuem taxas de sensibilidades mais altas e uma dispersão menor. Algo similar pode ser visualizado nas especificidades obtidas, pois esses métodos propostos têm valores altos (apesar de não serem os mais altos) e possuem as menores dispersões.

Figura 75 – Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 03.



Fonte: Elaborado pelo Autor.

Ao analisar as médias geométricas exibidas na Figura 76, é constatado que os métodos propostos tem resultados superiores aos clássicos, em ambas as abordagens. Porém, é necessário destacar que fazer o uso da abordagem intracanal é, em geral, melhor do que pela intercanais. Novamente, foi verificado que a parametrização com maior desempenho é a que se utiliza da esparsificação pela entropia de Renyi em conjunto da matriz de kernel RBF.

Como última análise para o presente caso, fez-se o uso da segmentação pela hipótese

Figura 76 – **CHB24**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 03.

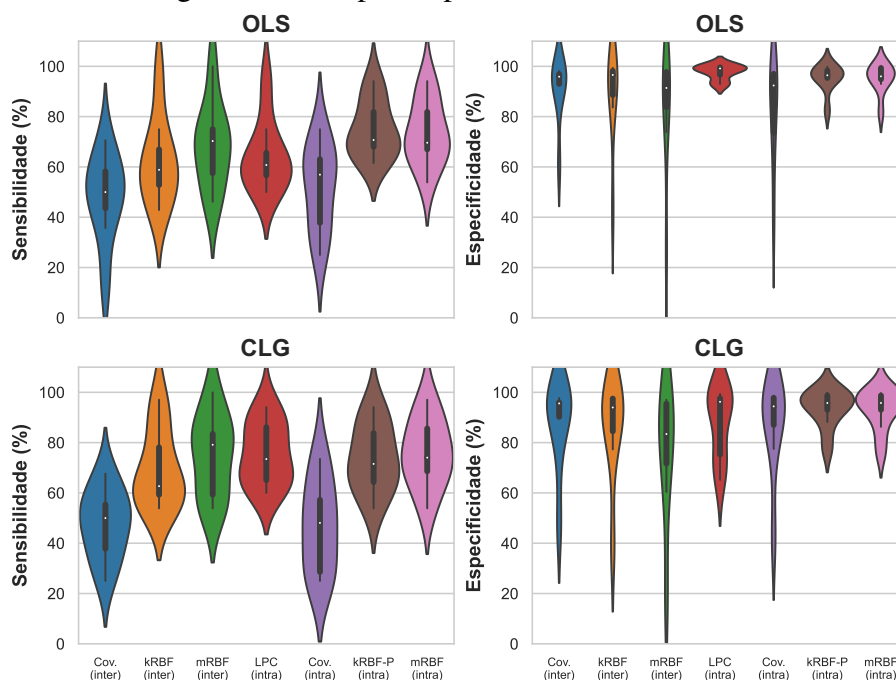


Fonte: Elaborado pelo Autor.

04. Com isto a aplicação dos classificadores OLS e CLG, produziram as taxas de sensibilidades e especificidades exibidas através da Figura 77. Pode-se verificar nesta figura, que no classificador OLS, as sensibilidades obtidas pelos métodos clássicos de parametrização, foram geralmente, inferiores aos métodos propostos pelo presente trabalho. Porém, com o uso do CLG, a sensibilidade do LPC obteve resultados similares aos dos métodos propostos. Analisando as especificidades obtidas pelo classificador OLS, verifica-se que o método em destaque é o LPC e em sequência estão os métodos propostos que utilizam a abordagem intracanal. Tal ocorrido se difere ao analisar as especificidades obtidas pelo CLG, já que o LPC possui uma inferioridade aos métodos propostos que fazem uso da abordagem intracanal.

Ao sumarizar os resultados obtidos pelas médias geométricas, é possível visualizar através da Figura 78 que também para esta segmentação os melhores resultados foram obtidos através dos métodos kRBF-P (intra) e mRBF (intra), que possuem valores respectivamente de 87,81% e 87,14%. Em sequência, encontra-se o método LPC que obteve valor de 84,88%. Os demais métodos propostos, apesar de terem taxas inferiores, ainda assim obtiveram valores superiores aos métodos de covariância (intra e inter). Ressalta-se que para esta segmentação a característica dos resultados obtidos pelos dois classificadores é similar, porém, o CLG tem resultados com magnitude levemente inferior. Por fim, disponibiliza-se o Quadro 9 presente no

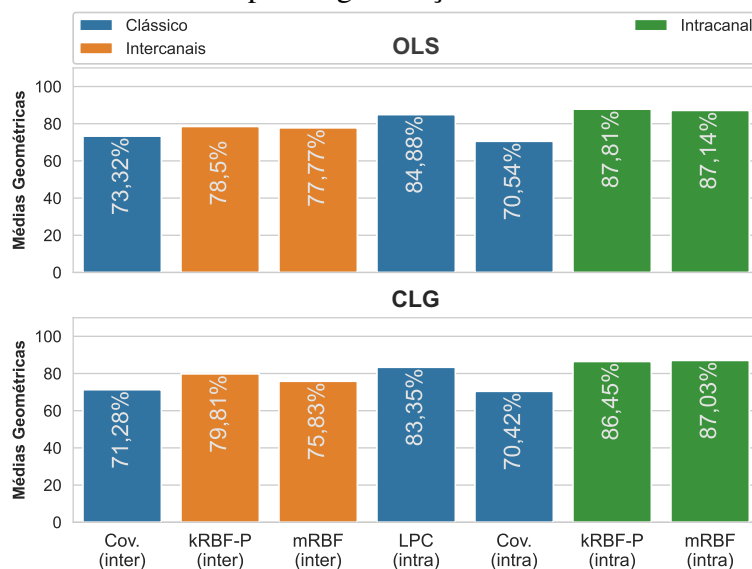
Figura 77 – **CHB24**: Comparativo entre sensibilidades e especificidades obtidas ao segmentar sinal pela hipótese 04.



Fonte: Elaborado pelo Autor.

Apêndice E que exibe os valores médios das sensibilidades, especificidades e acurácias obtidas em cada uma das quatro hipóteses de segmentação.

Figura 78 – **CHB24**: Comparativo das Médias Geométricas entre OLS e CLG para segmentação 04.

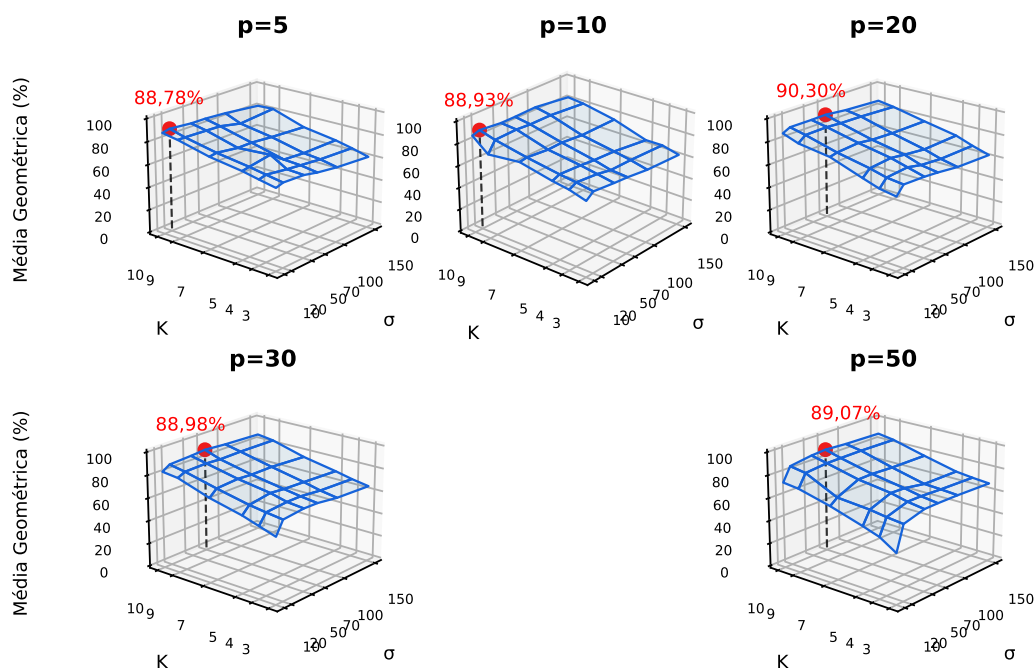


Fonte: Elaborado pelo Autor.

6.2 Resultados para o segundo conjunto de dados

Com relação aos resultados obtidos pelos sinais no segundo conjuntos de dados, inicialmente fez-se a exploração dos hiperparâmetros dos métodos utilizados no presente trabalho. Esta exploração foi realizada similarmente ao primeiro conjunto de dados. Na abordagem intracanal, faz-se necessária a definição do valor de p , a quantidade k ou m de vetores relevantes que compõe o dicionário e o parâmetro da função de kernel (σ ou d). Portanto, a primeira exploração foi realizada através da abordagem intracanal, pela esparsificação através do k -Médias e kernel RBF, como pode ser visualizado através da Figura 79.

Figura 79 – Exploração k Kernel RBF intracanal.

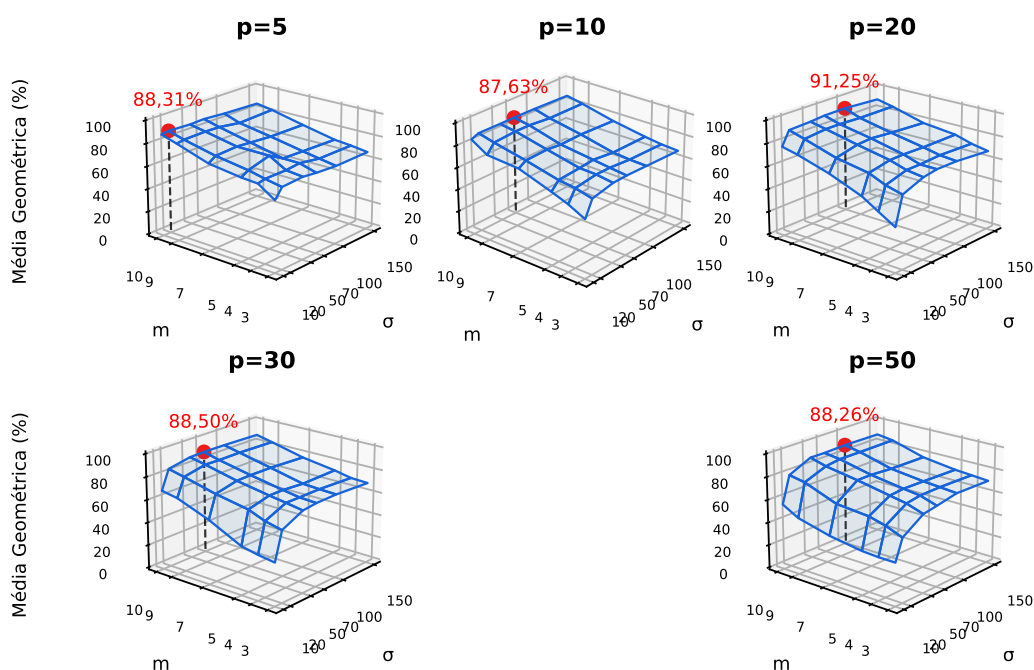


Fonte: Elaborado pelo Autor.

É possível verificar que a maior das médias geométricas obtidas, é encontrada quando $p = 20$, $k = 10$ e $\sigma = 100$. Vale ressaltar que a quantidade k para este conjunto de dados tem um impacto diferente do primeiro. Os sinais presentes no segundo conjunto de dados foram captados através de um equipamento que dispõe de 64 canais. Portanto, a exploração em k foi realizada até o valor 10, pois este gera um vetor de atributos por época de ordem $3520 + 1 \times 1$, limitando assim o custo computacional requisitado para esta etapa de exploração.

Em sequência, fez-se a exploração dos hiperparâmetros utilizados na parametrização que se utiliza da esparsificação através do critério de maximização da entropia de Renyi em conjunto com a matriz de kernel RBF. Como pode-se visualizar através da Figura 80 a taxa de média geométrica com maior valor (91,25%), foi encontrada quando se definiram os valores $p = 20$, $m = 10$ e $\sigma = 100$.

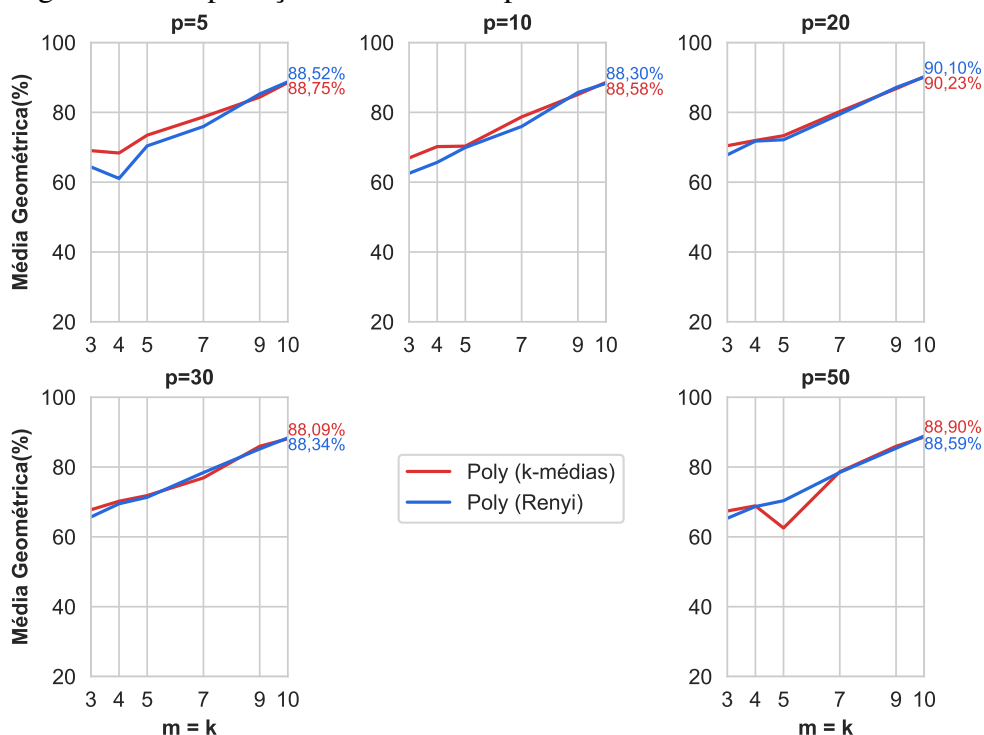
Figura 80 – Exploração m Kernel RBF intracanal.



Fonte: Elaborado pelo Autor.

Para explorar os hiperparâmetros ao utilizar o kernel polinomial, fez-se apenas uma limitação na ordem do polinômio. Isto foi realizado, pois se verificou um padrão de queda no desempenho ao utilizar ordens acima de um. Desta maneira, para construir a Figura 81 que ilustra a exploração, foram escolhidos os valores de $p = \{5, 10, 20, 30, 50\}$ e $m = k = \{3, 4, 5, 7, 9, 10\}$. Pode-se verificar que as maiores médias geométricas foram encontradas quando $p = 20$ e $m = k = 10$, para ambos os métodos de esparsificação. Ressalta-se que os k vetores mais próximos aos protótipos foram utilizados para montar o dicionário de vetores pelo k -Médias.

Com os hiperparâmetros definidos, realizaram-se 100 rodadas de treino e teste, embaralhando os vetores do conjunto de dados e dividindo 80% dos dados para treino e 20% para teste. Inicialmente esse processo foi conduzido através do classificador OLS, e os resultados

Figura 81 – Exploração $m = k$ Kernel polinomial intracanal.

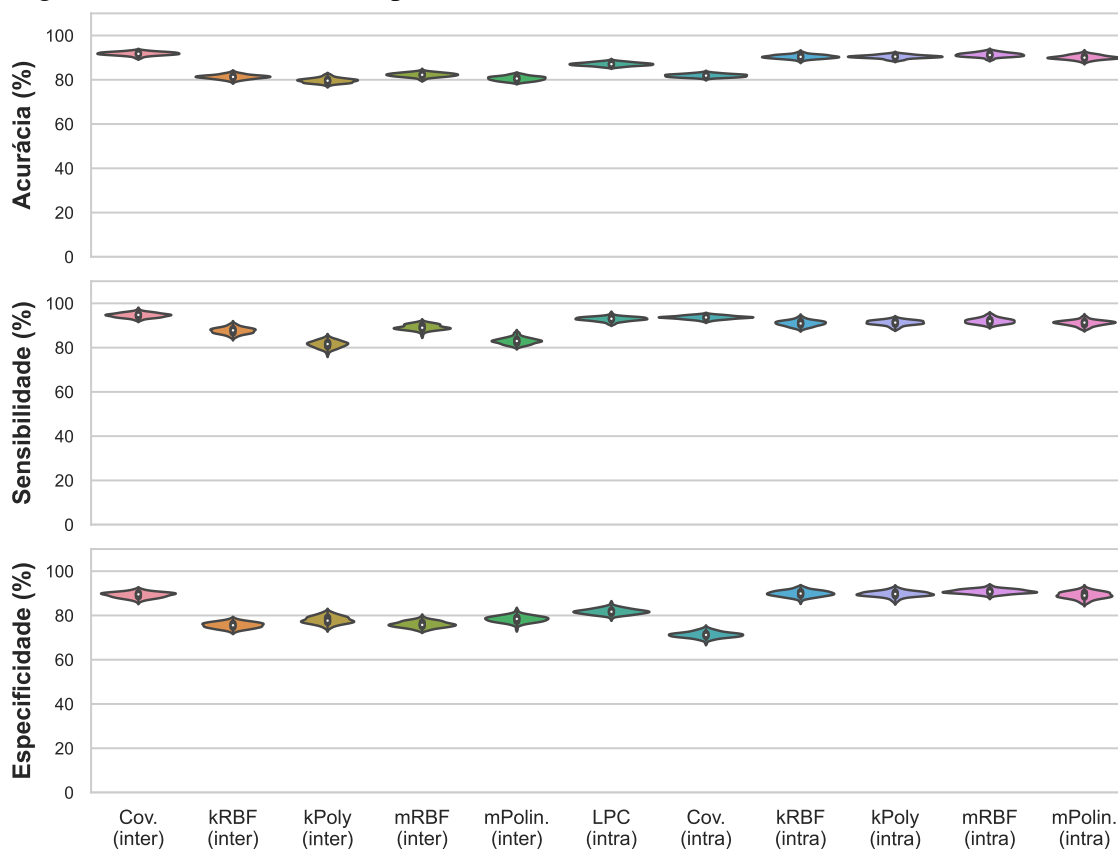
Fonte: Elaborado pelo Autor.

de sensibilidade, acurácia e especificidade obtidos, são exibidos através dos *violin plots* presentes na Figura 82. Pode-se perceber que como este conjunto de dados possui sinais mais controlados, as taxas obtidas não possuem a característica de dispersão observada nas análises do primeiro conjunto de dados. Verifica-se também que o grupo de métodos com melhores taxas, possuem valores próximo de 90%.

Analisando as sensibilidades e especificidades obtidas, percebe-se que o método de covariância intercanais, possui superioridade aos demais métodos. Tal ocorrido é de ser esperado por se tratar de um método espacial. Analisando os demais resultados, percebe-se que há um comportamento de desempenho promissor para os métodos propostos que se utilizaram da abordagem intracanal, com ênfase no método mRBF, que possui uma leve vantagem aos demais. Pode-se ainda perceber que os métodos propostos, aplicados na abordagem intercanais possuem desempenhos de sensibilidades e especificidades visivelmente inferiores aos que se utilizam da abordagem intracanal. Todavia, a utilização do método de covariância intracanal, não conseguiu desempenhar resultados tão bons quanto os demais métodos.

De modo a exibir os valores médios das 100 rodadas do algoritmo OLS, compôs-se o Tabela 3, que além destas informações, carrega o valor de médias geométricas computadas a partir das taxas médias de acurácia, sensibilidade e especificidade. Pode-se constatar ao verificar

Figura 82 – Métricas obtidas pelo classificador OLS.



Fonte: Elaborado pelo Autor.

estes valores que o método de covariância para o presente conjunto de dados se mostra como o que possui o melhor dos desempenhos com uma média geométrica de 91,86%. Ressalta-se que os métodos propostos pelo presente trabalho que utilizam da abordagem intracanal, são os que mais se aproximam dos valores obtidos pelo método da covariância. Dentre estes métodos, pode-se destacar o mRBF intracanal que possui média geométrica com valor de 91,43%. Os demais métodos da abordagem intracanal possuem médias geométricas acima de 90% e para alguns casos, a especificidade obtida é superior ao método de covariância. É importante destacar os desempenhos obtidos pelos métodos que usaram a matriz de kernel polinomial, que para o segundo conjunto de dados, tiveram resultados tão bons quanto o uso da matriz de kernel RBF.

Quando se aplica o algoritmo do CLG nas matrizes parametrizadas, os resultados têm comportamento similar como pode ser visualizado através da Figura 83. Nesta são exibidos os *violin plot* referentes as taxas de sensibilidade, especificidade e acurácia de todos os métodos. Pode-se perceber que os melhores métodos ainda são os que se utilizam da matriz de kernel numa abordagem intracanal, o da matriz de covariância intercanais e o LPC. Os demais métodos se encontram numa faixa de magnitude inferior, porém, destaca-se que para todos os métodos

Tabela 3 – Resultados médios ao utilizar o OLS.

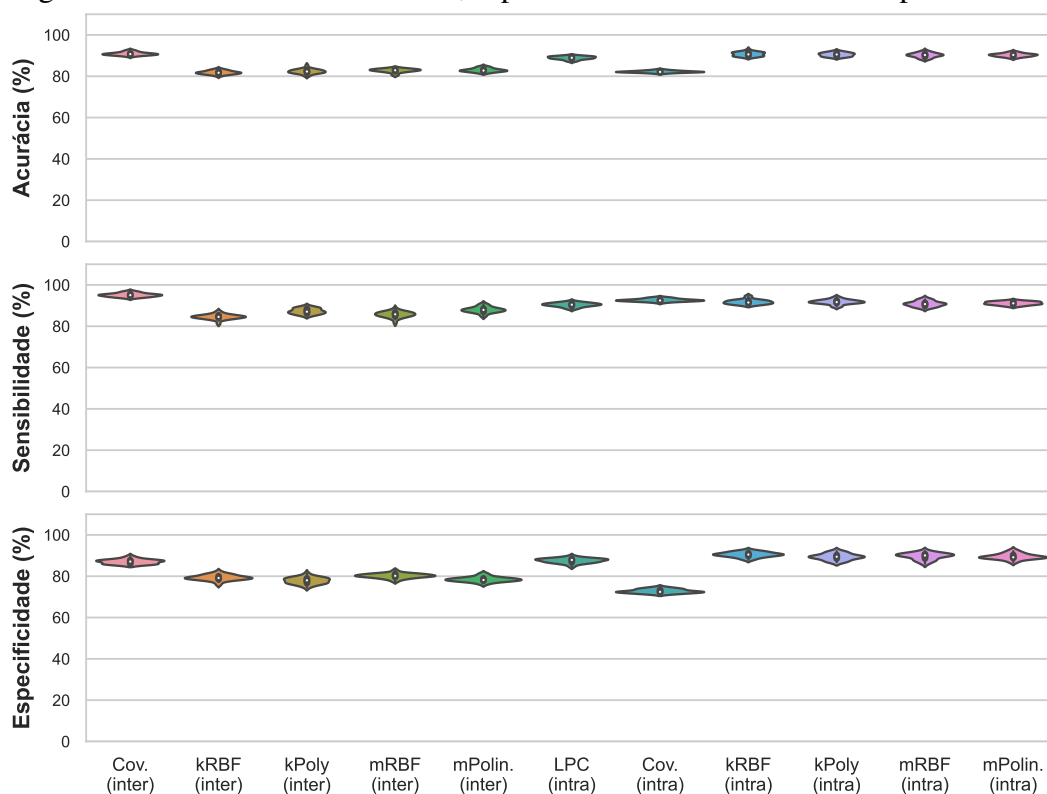
	Média Geom.	Acurácia	Especificidade	Sensibilidade
Covariância (inter)	91,86%	91,79%	89,25%	94,61%
<i>k</i> –Médias RBF (inter)	81,34%	81,28%	75,56%	87,63%
<i>k</i> –Médias Polin. (inter)	79,66%	79,61%	77,93%	81,48%
<i>m</i> –Renyi RBF (inter)	82,25%	82,20%	75,92%	89,17%
<i>m</i> –Renyi Polin. (inter)	80,61%	80,55%	78,37%	82,97%
LPC	87,12%	87,04%	81,73%	92,93%
Covariância (intra)	81,73%	81,86%	71,25%	93,62%
<i>k</i> –Médias RBF (intra)	90,40%	90,39%	89,87%	90,96%
<i>k</i> –Médias Polin. (intra)	90,43%	90,40%	89,65%	91,23%
<i>m</i> –Renyi RBF (intra)	91,34%	91,25%	90,78%	91,98%
<i>m</i> –Renyi Polin. (intra)	90,11%	90,08%	89,15%	91,11%

Fonte: Elaborado pelo Autor.

aplicados, ainda foi possível visualizar uma baixa dispersão das taxas obtidas.

A fim a quantificar os desempenhos obtidos ao utilizar o CLG, foi realizada a construção do Tabela 4, que exhibe as médias de cada uma das taxas presentes nos *violin plots*, bem como a média geométrica computada. Pode-se observar que neste caso, o método com maior valor de média geométrica (90,96%) é o proposto pelo presente trabalho, que utiliza a esparsificação pelo *k*–Médias em conjunto da matriz de kernel RBF através da abordagem intracanal. Em sequência, encontra-se o método de covariância, que possui uma média geométrica de 90,94%. Destacam-se também os demais métodos propostos pelo presente trabalho que se utilizam da abordagem intracanal. Neste caso, faz-se necessário também descrever que o kernel polinomial obteve resultados tão bons quanto os pela matriz de kernel RBF. Pode-se ainda afirmar que por conta da qualidade dos sinais disponíveis no presente conjunto de dados, as médias geométricas obtidas são todas acima de 80%. Todavia, faz-se necessário destacar que os métodos propostos aplicados pela abordagem intercanais, possuem valores de média geométrica inferiores aos demais métodos testados.

Figura 83 – Taxas de sensibilidade, especificidade e acurácia obtidas pelo CLG.



Fonte: Elaborado pelo Autor.

Tabela 4 – Resultados médios ao utilizar CLG.

	Média Geom.	Acurácia	Especificidade	Sensibilidade
Covariância (inter)	90,94%	90,86%	87,04%	95,10%
<i>k</i> -Médias RBF (inter)	81,84%	81,78%	79,29%	84,54%
<i>k</i> -Médias Polin. (inter)	82,38%	82,31%	77,84%	87,26%
<i>m</i> -Renyi RBF (inter)	82,83%	82,77%	80,15%	85,66%
<i>m</i> -Renyi Polin. (inter)	83,01%	82,94%	78,45%	87,90%
LPC	88,91%	88,87%	87,53%	90,36%
Covariância (intra)	82,12%	82,18%	72,76%	92,62%
<i>k</i> -Médias RBF (intra)	90,96%	90,69%	90,44%	91,74%
<i>k</i> -Médias Polin. (intra)	90,52%	90,48%	89,35%	91,74%
<i>m</i> -Renyi RBF (intra)	90,21%	90,19%	89,62%	90,83%
<i>m</i> -Renyi Polin. (intra)	90,28%	90,25%	89,49%	91,10%

Fonte: Elaborado pelo Autor.

6.3 Avaliação de custo computacional

A presente seção trata da avaliação dos custos computacionais dos métodos de parametrização utilizados no trabalho. Destaca-se que nesta etapa excluiu-se a análise de custo computacional dos métodos que fazem uso da esparsificação pelo k -Médias ao selecionar os k vetores protótipos. Ou seja, os métodos selecionados para construir os gráficos e tabelas da presente seção, são os mesmo utilizados para exibir os resultados do segundo conjunto de dados. Os algoritmos foram executados em um computador Intel(R) Core(TM) i3-4160 CPU @ 3.60GHz 3.60 GHz que possui 16,0 GB.

Os resultados são divididos entre os métodos aplicados no primeiro, e no segundo conjunto de dados. Isto foi realizado, pois, os hiperparâmetros se diferenciam para cada conjunto e dependendo do caso, ocasiona em diferentes tamanhos de vetores atributos que por consequência faz com que seus métodos desempenhem em tempos diferentes.

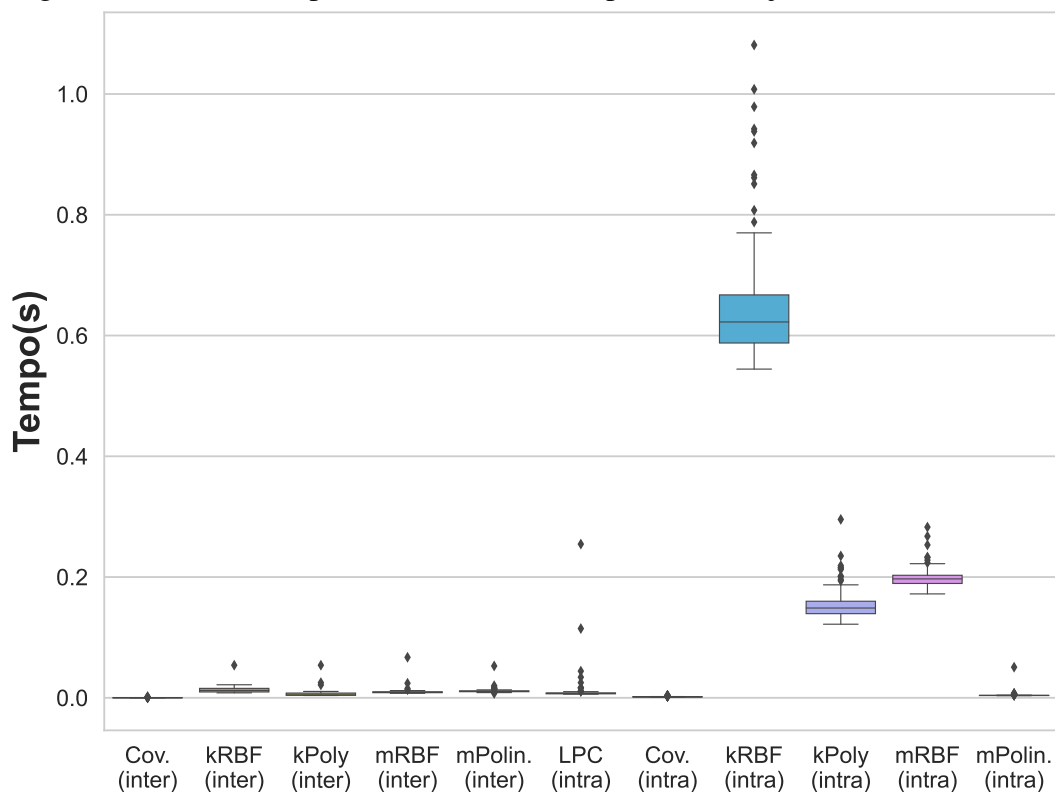
Pode-se perceber através da Figura 84, que exibe os *boxplots* referentes aos custos computacionais observados ao aplicar os métodos no primeiro conjunto de dados, que os clássicos desempenham suas tarefas em um baixo tempo. Todavia, os métodos pela análise intracanal propostos pelo presente trabalho, tem um elevado custo computacional. Com destaque para os métodos que utilizam a esparsificação pelo k -Médias.

Em contrapartida, os custos computacionais observados na aplicação do segundo conjunto de dados, são diferentes. Isto pode ser constatado ao visualizar o *boxplot* na Figura 85. Nesta, o que se difere da primeira é que, em geral, os custos são elevados, devido a dois fatores: Quantidade de canais; Hiperparâmetros selecionados. Nesta análise ainda pode-se dizer também que os métodos clássicos têm baixo custo computacional, em comparação com os melhores métodos propostos. Novamente o método que utiliza a esparsificação pelo k -Médias tem custo superior aos demais.

6.4 Considerações Finais

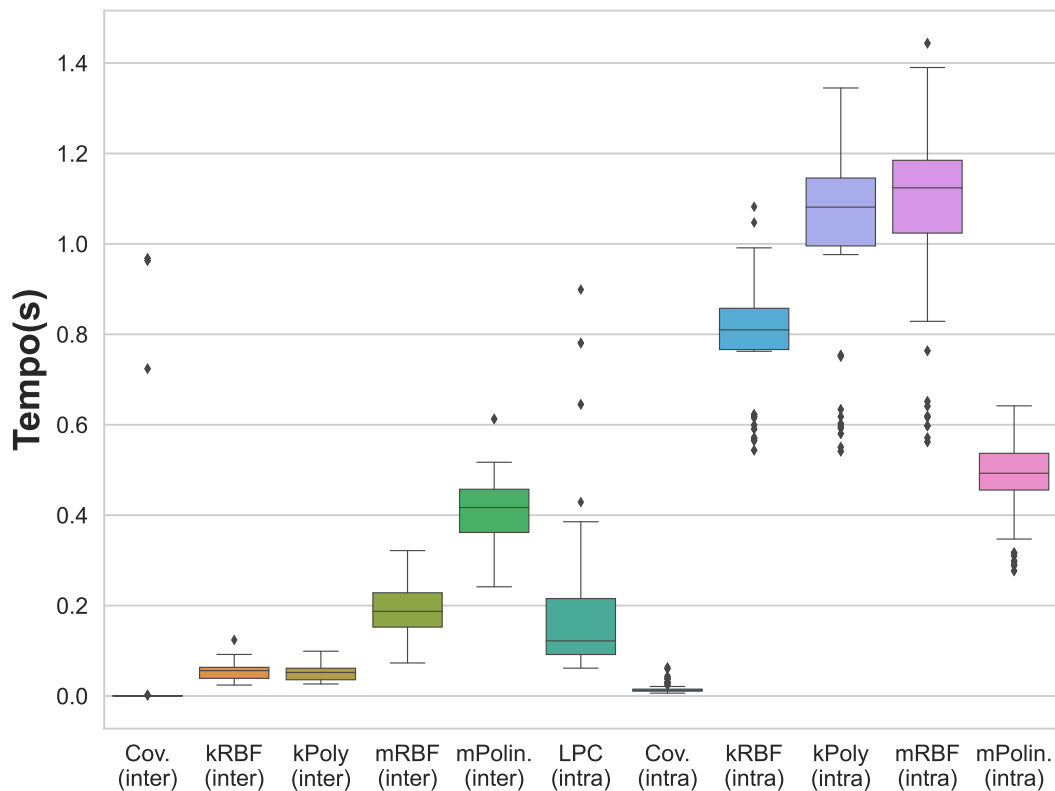
Neste capítulo, foram expostos os resultados obtidos ao aplicar os métodos clássicos e propostos de parametrização dos sinais de EEG. Em uma análise inicial, buscou-se identificar os parâmetros que proporcionam um aumento do desempenho dos modelos na etapa de classificação. Em sequência, considerando apenas o primeiro conjunto de dados, fez-se a utilização dos parâmetros encontrados para diferentes casos presentes neste conjunto, compondo assim os

Figura 84 – Custo computacional métodos no primeiro conjunto de dados.



Fonte: Elaborado pelo Autor.

Figura 85 – Custo computacional métodos no segundo conjunto de dados.



Fonte: Elaborado pelo Autor.

resultados individuais para cada paciente analisado. Nessa análise extensiva foi possível verificar que os métodos propostos pela presente dissertação são promissores, pois, tem desempenhos similares ou superiores aos métodos clássicos. Após a análise realizada para o primeiro conjunto de dados, o procedimento é repetido para o segundo, em que foi possível verificar também que os métodos propostos não lineares tiveram desempenhos similares aos clássicos lineares. O capítulo se encerra pela exposição da análise de custo computacional associado a cada variação dos métodos de parametrização utilizados no presente trabalho. Nesta análise é possível verificar que as abordagens intracanal que fazem o uso da estimação das matrizes de kernel, tem um custo por época maior que os demais métodos.

7 CONCLUSÕES E TRABALHOS FUTUROS

Nesta dissertação, propuseram-se novas abordagens para parametrização de sinais de EEG via matrizes de kernel, em que foram usadas estratégias intercanais e intracanal para construção do vetor de atributos. Para isso, realizou-se um levantamento bibliográfico sobre as características de sinais EEG, extração de atributos, conceitos de kernel, teoria de esparsificação de matrizes e algoritmos de classificação. Os experimentos conduzidos, foram realizados em dois diferentes conjuntos de dados, o que envolveu um amplo processo investigativo nos hiperparâmetros dos métodos utilizados. Em seguida, procurou-se avaliar os métodos propostos através de um levantamento comparativo entre os desempenhos dos métodos propostos usando algoritmos de classificação lineares chamados.

A partir dos resultados obtidos, pode-se chegar às conclusões listadas abaixo.

1. Os métodos clássicos de covariância intercanais e o LPC, conseguem desempenhar bem a tarefa de parametrização.
2. O método clássico de covariância intracanal, teve desempenhos inferiores, portanto, tem desempenhos insatisfatórios.
3. Com respeito aos métodos propostos, a utilização das análises intracanal possui geralmente, um desempenho superior à abordagem intercanais.
4. O método proposto que se utiliza da esparsificação através do método k -Médias em conjunto da matriz de kernel RBF e pela abordagem intracanal, é promissor e em alguns casos se obteve um desempenho superior aos métodos clássicos.
5. O método proposto que se utiliza da esparsificação através do critério de maximização da entropia de Renyi em conjunto da matriz de kernel RBF e pela abordagem intracanal, tem resultados promissores. Entre os métodos propostos, este se destaca dos demais, já em diversos casos, este possui o maior desempenho observado.
6. Observou-se que a utilização do kernel polinomial quando aplicado em modo intercanal (independente da esparsificação), se obteve resultados não promissores. Dessa maneira, se encontram num patamar de inferioridade dos métodos propostos e clássicos.
7. Pelos resultados obtidos no segundo conjunto de dados, pôde-se perceber que o kernel polinomial pela abordagem intracanal, teve desempenhos tão bons quanto aos métodos clássicos ou os que se utilizam do kernel RBF.
8. Ao realizar as segmentações presentes no primeiro conjunto de dados, foi possível verificar que em diversos casos, os resultados obtidos pela hipótese 2, eram inferiores às demais

segmentações. Destaca-se ainda que com o uso dessas segmentações, foi possível confirmar o bom desempenho de alguns métodos propostos.

9. Com a utilização dos classificadores presentes no trabalho, pode-se verificar que o método OLS tem geralmente resultados melhores que o CLG. Porém, na aplicação do CLG, foi percebido uma melhora nos piores desempenhos obtidos pelo OLS. Enquadram-se nesta categoria os métodos que utilizam do kernel polinomial.
10. Pode-se notar que o custo computacional dos métodos propostos pelo presente trabalho, são superiores aos métodos clássicos.

7.1 Trabalhos futuros

As seguintes linhas de investigação podem ser trabalhadas como uma extensão do estudo realizado nesta dissertação.

- a) Participar do processo de aquisições de sinais EEG em ambiente clínico, de modo a construir um conjunto de dados conhecido.
- b) Realizar variações dos métodos propostos, através da utilização de outras funções de kernel conhecidas.
- c) Propor uma construção de nova matriz de kernel para realizar a parametrização dos sinais EEG.
- d) Propor uma nova abordagem para construir o dicionário de vetores relevantes.
- e) Conduzir novos experimentos na etapa de classificação, através escolha de algoritmos não lineares como, por exemplo, redes neurais artificiais.

REFERÊNCIAS

- ACHARYA, J. N.; HANI, A. J.; CHEEK, J.; THIRUMALA, P.; TSUCHIDA, T. N. American clinical neurophysiology society guideline 2: Guidelines for standard electrode position nomenclature. **The Neurodiagnostic Journal**, Informa UK Limited, v. 56, n. 4, p. 245–252, out. 2016. Disponível em: <https://doi.org/10.1080/21646821.2016.1245558>.
- ACHARYA, U. R.; BHAT, S.; FAUST, O.; ADELI, H.; CHUA, E. C.-P.; LIM, W. J. E.; KOH, J. E. W. Nonlinear dynamics measures for automated EEG-based sleep stage detection. **European Neurology**, S. Karger AG, v. 74, n. 5-6, p. 268–287, dez. 2015. Disponível em: <https://doi.org/10.1159/000441975>.
- ALÇIN, Ö. F.; SIULY, S.; BAJAJ, V.; GUO, Y.; ŞENGU, A.; ZHANG, Y. *et al.* Multi-category EEG signal classification developing time-frequency texture features based fisher vector encoding method. **Neurocomputing**, Elsevier BV, v. 218, p. 251–258, dez. 2016. Disponível em: <https://doi.org/10.1016/j.neucom.2016.08.050>.
- ALFARAS, M.; SORIANO, M. C.; ORTÍN, S. A fast machine learning model for ECG-based heartbeat classification and arrhythmia detection. **Frontiers in Physics**, Frontiers Media SA, v. 7, jul. 2019. Disponível em: <https://doi.org/10.3389/fphy.2019.00103>.
- BAIG, M. Z.; ASLAM, N.; SHUM, H. P. H. Filtering techniques for channel selection in motor imagery EEG applications: a survey. **Artificial Intelligence Review**, Springer Science and Business Media LLC, v. 53, n. 2, p. 1207–1232, fev. 2019. Disponível em: <https://doi.org/10.1007/s10462-019-09694-8>.
- BARACHANT, A.; BONNET, S.; CONGEDO, M.; JUTTEN, C. Classification of covariance matrices using a riemannian-based kernel for BCI applications. **Neurocomputing**, Elsevier BV, v. 112, p. 172–178, jul. 2013. Disponível em: <https://doi.org/10.1016/j.neucom.2012.12.039>.
- BAYS, B. C.; VISSCHER, K. M.; DANTEC, C. C. L.; SEITZ, A. R. Alpha-band EEG activity in perceptual learning. **Journal of Vision**, Association for Research in Vision and Ophthalmology (ARVO), v. 15, n. 10, p. 7, set. 2015. Disponível em: <https://doi.org/10.1167/15.10.7>.
- BEGLEITER, H. Neurodynamics laboratory. **State University of New York Health Center at Brooklyn**, 1999.
- BERGER, H. Über das elektroencephalogramm des menschen. **Archiv für psychiatrie und nervenkrankheiten**, v. 87, n. 1, p. 527–570, 1929.
- BIASIUCCI, A.; FRANCESCHIELLO, B.; MURRAY, M. M. Electroencephalography. **Current Biology**, Elsevier, v. 29, n. 3, p. R80–R85, 2019.
- BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: **Proceedings of the fifth annual workshop on Computational learning theory - COLT '92**. ACM Press, 1992. Disponível em: <https://doi.org/10.1145/130385.130401>.
- BÖTTCHER, A. **Introduction to large truncated Toeplitz matrices**. New York: Springer, 1999. ISBN 9780387985701.
- BRIENZA, M.; DAVASSI, C.; MECARELLI, O. Artifacts. In: **Clinical Electroencephalography**. [S. l.]: Springer, 2019. p. 109–130.

CARDOSO, A. S. V. Instrumentação e metodologias de medição de biopotenciais. Universidade Federal de Minas Gerais, 2010.

DELORME, A.; MAKEIG, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. **Journal of Neuroscience Methods**, Elsevier BV, v. 134, n. 1, p. 9–21, mar. 2004. Disponível em: <https://doi.org/10.1016/j.jneumeth.2003.10.009>.

DORA, C.; BISWAL, P. K. Robust ECG artifact removal from EEG using continuous wavelet transformation and linear regression. In: **2016 International Conference on Signal Processing and Communications (SPCOM)**. IEEE, 2016. Disponível em: <https://doi.org/10.1109/spcom.2016.7746620>.

DOUPE, P.; FAGHMOUS, J.; BASU, S. Machine learning for health services researchers. **Value in Health**, Elsevier BV, v. 22, n. 7, p. 808–815, jul. 2019. Disponível em: <https://doi.org/10.1016/j.jval.2019.02.012>.

ELSEVIER, C. L. Gray's anatomy. the anatomical basis of clinical practice. **American Journal of Neuroradiology**, v. 26, n. 10, 2005.

FATTINGER, S.; KURTH, S.; RINGLI, M.; JENNI, O. G.; HUBER, R. Theta waves in children's waking electroencephalogram resemble local aspects of sleep during wakefulness. **Scientific Reports**, Springer Science and Business Media LLC, v. 7, n. 1, set. 2017. Disponível em: <https://doi.org/10.1038/s41598-017-11577-3>.

GE, S.; YANG, Q.; WANG, R.; LIN, P.; GAO, J.; LENG, Y.; YANG, Y.; WANG, H. A brain-computer interface based on a few-channel EEG-fNIRS bimodal system. **IEEE Access**, Institute of Electrical and Electronics Engineers (IEEE), v. 5, p. 208–218, 2017. Disponível em: <https://doi.org/10.1109/access.2016.2637409>.

GOLDENBERG, S. L.; NIR, G.; SALCUDEAN, S. E. A new era: artificial intelligence and machine learning in prostate cancer. **Nature Reviews Urology**, Springer Science and Business Media LLC, v. 16, n. 7, p. 391–403, maio 2019. Disponível em: <https://doi.org/10.1038/s41585-019-0193-3>.

GOLOVKO, V.; BEZOBRAZOVA, S. Neural networks for chaotic signal processing: Application to the electroencephalogram analysis for epilepsy detection. 07 2021.

HAN, M.; LIANG, Z.; LI, D. Sparse kernel density estimations and its application in variable selection based on quadratic renyi entropy. **Neurocomputing**, Elsevier BV, v. 74, n. 10, p. 1664–1672, maio 2011. Disponível em: <https://doi.org/10.1016/j.neucom.2011.01.022>.

HENRY, J. C. Electroencephalography: basic principles, clinical applications, and related fields. **Neurology**, AAN Enterprises, v. 67, n. 11, p. 2092–2092, 2006.

HOMAN, R. W. The 10-20 electrode system and cerebral location. **American Journal of EEG Technology**, Taylor & Francis, v. 28, n. 4, p. 269–279, 1988.

HOSSEINI, M.-P.; HOSSEINI, A.; AHI, K. A review on machine learning for EEG signal processing in bioengineering. **IEEE Reviews in Biomedical Engineering**, Institute of Electrical and Electronics Engineers (IEEE), v. 14, p. 204–218, 2021. Disponível em: <https://doi.org/10.1109/rbme.2020.2969915>.

HU, K. Become competent within one day in generating boxplots and violin plots for a novice without prior r experience. **Methods and Protocols**, MDPI AG, v. 3, n. 4, p. 64, set. 2020. Disponível em: <https://doi.org/10.3390/mps3040064>.

HYNDMAN, R. J.; SHANG, H. L. Rainbow plots, bagplots, and boxplots for functional data. **Journal of Computational and Graphical Statistics**, Informa UK Limited, v. 19, n. 1, p. 29–45, jan. 2010. Disponível em: <https://doi.org/10.1198/jcgs.2009.08158>.

JAFARIFARMAND, A.; BADAMCHIZADEH, M. A. Artifacts removal in EEG signal using a new neural network enhanced adaptive filter. **Neurocomputing**, Elsevier BV, v. 103, p. 222–231, mar. 2013. Disponível em: <https://doi.org/10.1016/j.neucom.2012.09.024>.

JAMBUKIA, S. H.; DABHI, V. K.; PRAJAPATI, H. B. Classification of ECG signals using machine learning techniques: A survey. In: **2015 International Conference on Advances in Computer Engineering and Applications**. IEEE, 2015. Disponível em: <https://doi.org/10.1109/icacea.2015.7164783>.

JAMES, G. **An introduction to statistical learning : with applications in R**. New York: Springer, 2021. ISBN 9781071614174.

JASPER, H. H. The ten-twenty electrode system of the international federation. **Electroencephalogr. Clin. Neurophysiol.**, v. 10, p. 370–375, 1958.

JONES, S. L.; SVITKINA, T. M. Axon initial segment cytoskeleton: Architecture, development, and role in neuron polarity. **Neural Plasticity**, Hindawi Limited, v. 2016, p. 1–19, 2016. Disponível em: <https://doi.org/10.1155/2016/6808293>.

KANDEL, E. R.; SCHWARTZ, J. H.; JESSELL, T. M.; SIEGELBAUM, S.; HUDSPETH, A. J.; MACK, S. **Principles of neural science**. [S. l.]: McGraw-hill New York, 2000. v. 4.

KANE, N.; ACHARYA, J.; BENICZKY, S.; CABOCLO, L.; FINNIGAN, S.; KAPLAN, P. W.; SHIBASAKI, H.; PRESSLER, R.; PUTTEN, M. J. van. A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings. revision 2017. **Clinical Neurophysiology Practice**, Elsevier BV, v. 2, p. 170–185, 2017. Disponível em: <https://doi.org/10.1016/j.cnp.2017.07.002>.

KAPPEL, S. L.; LOONEY, D.; MANDIC, D. P.; KIDMOSE, P. Physiological artifacts in scalp EEG and ear-EEG. **BioMedical Engineering OnLine**, Springer Science and Business Media LLC, v. 16, n. 1, ago. 2017. Disponível em: <https://doi.org/10.1186/s12938-017-0391-2>.

KHOUMA, O.; DIOP, I.; FALL, P. A.; NDIAYE, M. L.; FARSSI, S. M.; OUSSAMATOU, A. M.; DIOUF, B. Novel classification method of spikes morphology in EEG signal using machine learning. **Procedia Computer Science**, Elsevier BV, v. 148, p. 70–79, 2019. Disponível em: <https://doi.org/10.1016/j.procs.2019.01.010>.

KIRKCALDIE, M. T.; COLLINS, J. M. The axon as a physical structure in health and acute trauma. **Journal of Chemical Neuroanatomy**, Elsevier BV, v. 76, p. 9–18, out. 2016. Disponível em: <https://doi.org/10.1016/j.jchemneu.2016.05.006>.

KLEM, G. H. The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology. **Electroencephalogr. Clin. Neurophysiol. Suppl.**, v. 52, p. 3–6, 1999.

KOCHETUROV, A.; PARDALOS, P. M.; KARAKITSIOU, A. Massive datasets and machine learning for computational biomedicine: trends and challenges. **Annals of Operations Research**, Springer Science and Business Media LLC, v. 276, n. 1-2, p. 5–34, maio 2018. Disponível em: <https://doi.org/10.1007/s10479-018-2891-2>.

KOLKHORST, H.; VEIT, J.; BURGARD, W.; TANGERMANN, M. A robust screen-free brain-computer interface for robotic object selection. **Frontiers in Robotics and AI**, Frontiers Media SA, v. 7, mar. 2020. Disponível em: <https://doi.org/10.3389/frobt.2020.00038>.

KRASOULIS, A.; VIJAYAKUMAR, S.; NAZARPOUR, K. Multi-grip classification-based prosthesis control with two EMG-IMU sensors. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, Institute of Electrical and Electronics Engineers (IEEE), v. 28, n. 2, p. 508–518, fev. 2020. Disponível em: <https://doi.org/10.1109/tnsre.2019.2959243>.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, Springer Science and Business Media LLC, v. 521, n. 7553, p. 436–444, maio 2015. Disponível em: <https://doi.org/10.1038/nature14539>.

LONDON, M.; HÄUSSER, M. Dendritic computation. **Annu. Rev. Neurosci.**, Annual Reviews, v. 28, p. 503–532, 2005.

M., H. D.; CID, F. M.; OTÁROLA, J.; ROJAS, R.; ALARCÓN, O.; CAÑETE, L. EEG beta band frequency domain evaluation for assessing stress and anxiety in resting, eyes closed, basal conditions. **Procedia Computer Science**, Elsevier BV, v. 162, p. 974–981, 2019. Disponível em: <https://doi.org/10.1016/j.procs.2019.12.075>.

MASZCZYK, A.; DOBRAKOWSKI, P.; ŻAK, M.; GOZDOWSKI, P.; KRAWCZYK, M.; MAŁECKI, A.; STASTNY, P.; ZAJAC, T. Differences in motivation during the bench press movement with progressive loads using EEG analysis. **Biology of Sport**, Termedia Sp. z.o.o., v. 36, n. 4, p. 351–356, 2019. Disponível em: <https://doi.org/10.5114/biolsport.2019.88757>.

MCGWIRE, K. C.; FISHER, P. Spatially variable thematic accuracy: Beyond the confusion matrix. In: **Spatial Uncertainty in Ecology**. Springer New York, 2001. p. 308–329. Disponível em: https://doi.org/10.1007/978-1-4613-0209-4_14.

MECARELLI, O. Past, present and future of the eeg. In: **Clinical Electroencephalography**. [S. l.]: Springer, 2019. p. 3–8.

MILLIKEN, G. **Analysis of messy data**. Belmont, Calif: Lifetime Learning Publications, 1984. ISBN 9781584880837.

MO, H.; ZHAO, Y. Motor imagery electroencephalograph classification based on optimized support vector machine by magnetic bacteria optimization algorithm. **Neural Processing Letters**, Springer Science and Business Media LLC, v. 44, n. 1, p. 185–197, set. 2015. Disponível em: <https://doi.org/10.1007/s11063-015-9469-7>.

NEUMAN, M. R. *et al.* Biopotential electrodes. **Medical instrumentation: application and design**, Wiley New York, v. 4, p. 189–240, 1998.

NIEDERMEYER, E.; SILVA, F. L. D. **Electroencephalography–Basic principles, clinical applications, and related fields**. [S. l.]: Urban & Schwarzenberg,, 2020.

NIK, H. **Introduction to probability, statistics, and random processes**. Blue Bell, PA: Kappa Research, LLC, 2014. ISBN 9780990637202.

OLIVEIRA, A. M. S. **Extração de Características de sinais de eletroencefalograma usando matriz de covariância para identificação de crises epiléticas**. 2021. Trabalho de Conclusão de Curso (Bacharel em Engenharia de Computação), DETI (Departamento de Engenharia de Teleinformática), UFC (Universidade Federal do Ceará) Fortaleza, Brasil.

PISE, A. W.; REGE, P. P. Comparative analysis of various filtering techniques for denoising EEG signals. In: **2021 6th International Conference for Convergence in Technology (I2CT)**. IEEE, 2021. Disponível em: <https://doi.org/10.1109/i2ct51068.2021.9417984>.

QAYYUM, A.; QADIR, J.; BILAL, M.; AL-FUQAHA, A. Secure and robust machine learning for healthcare: A survey. **IEEE Reviews in Biomedical Engineering**, Institute of Electrical and Electronics Engineers (IEEE), v. 14, p. 156–180, 2021. Disponível em: <https://doi.org/10.1109/rbme.2020.3013489>.

RAMOS-AGUILAR, R.; OLVERA-LÓPEZ, J. A.; OLMOS-PINEDA, I.; SÁNCHEZ-URRIETA, S. Feature extraction from EEG spectrograms for epileptic seizure detection. **Pattern Recognition Letters**, Elsevier BV, v. 133, p. 202–209, maio 2020. Disponível em: <https://doi.org/10.1016/j.patrec.2020.03.006>.

RASHEED, K.; QAYYUM, A.; QADIR, J.; SIVATHAMBOO, S.; KWAN, P.; KUHLMANN, L.; O'BRIEN, T.; RAZI, A. Machine learning for predicting epileptic seizures using EEG signals: A review. **IEEE Reviews in Biomedical Engineering**, Institute of Electrical and Electronics Engineers (IEEE), v. 14, p. 139–155, 2021. Disponível em: <https://doi.org/10.1109/rbme.2020.3008792>.

REN, W.; HAN, M. Classification of EEG signals using hybrid feature extraction and ensemble extreme learning machine. **Neural Processing Letters**, Springer Science and Business Media LLC, v. 50, n. 2, p. 1281–1301, set. 2018. Disponível em: <https://doi.org/10.1007/s11063-018-9919-0>.

SAIRAMYA, N. J.; GEORGE, S. T.; SUBATHRA, M. S. P.; KUMAR, N. M. Computer-aided diagnosis of epilepsy based on the time-frequency texture descriptors of EEG signals using wavelet packet decomposition and artificial neural network. In: **Cognitive Informatics and Soft Computing**. Springer Singapore, 2018. p. 677–688. Disponível em: https://doi.org/10.1007/978-981-13-0617-4_64.

SANTOS, J. D. A.; BARRETO, G. A. Novel sparse LSSVR models in primal weight space for robust system identification with outliers. **Journal of Process Control**, Elsevier BV, v. 67, p. 129–140, jul. 2018. Disponível em: <https://doi.org/10.1016/j.jprocont.2017.04.001>.

SAVADKOOHI, M.; OLADUNNI, T.; THOMPSON, L. A machine learning approach to epileptic seizure prediction using electroencephalogram (EEG) signal. **Biocybernetics and Biomedical Engineering**, Elsevier BV, v. 40, n. 3, p. 1328–1341, jul. 2020. Disponível em: <https://doi.org/10.1016/j.bbe.2020.07.004>.

SCHÖLKOPF, B.; SMOLA, A. J.; BACH, F. *et al.* **Learning with kernels: support vector machines, regularization, optimization, and beyond**. [S. l.]: MIT press, 2002.

- SEECK, M.; KOESSLER, L.; BAST, T.; LEIJTEN, F.; MICHEL, C.; BAUMGARTNER, C.; HE, B.; BENICZKY, S. The standardized EEG electrode array of the IFCN. **Clinical Neurophysiology**, Elsevier BV, v. 128, n. 10, p. 2070–2077, out. 2017. Disponível em: <https://doi.org/10.1016/j.clinph.2017.06.254>.
- SHAD, E. H. T.; MOLINAS, M.; YTTERDAL, T. Impedance and noise of passive and active dry EEG electrodes: A review. **IEEE Sensors Journal**, Institute of Electrical and Electronics Engineers (IEEE), v. 20, n. 24, p. 14565–14577, dez. 2020. Disponível em: <https://doi.org/10.1109/jsen.2020.3012394>.
- SHARMA, R.; SIRCAR, P.; PACHORI, R. B. Automated focal EEG signal detection based on third order cumulant function. **Biomedical Signal Processing and Control**, Elsevier BV, v. 58, p. 101856, abr. 2020. Disponível em: <https://doi.org/10.1016/j.bspc.2020.101856>.
- SHAWE-TAYLOR, J.; CRISTIANINI, N. *et al.* **Kernel methods for pattern analysis**. [S. l.]: Cambridge university press, 2004.
- SHOEB, A. **CHB-MIT Scalp EEG Database**. physionet.org, 2010. Disponível em: <https://physionet.org/content/chbmit/>.
- SHOEB, A. H. **Application of machine learning to epileptic seizure onset detection and treatment**. Tese (Doutorado) – Massachusetts Institute of Technology, 2009.
- SHUMWAY, R. **Time series analysis and its applications : with R examples**. Cham, Switzerland: Springer, 2017. ISBN 9783319524511.
- SIJLY, S.; ZHANG, Y. Medical big data: Neurological diseases diagnosis through medical data analysis. **Data Science and Engineering**, Springer Science and Business Media LLC, v. 1, n. 2, p. 54–64, jun. 2016. Disponível em: <https://doi.org/10.1007/s41019-016-0011-3>.
- SNODGRASS, J. G.; VANDERWART, M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. **Journal of Experimental Psychology: Human Learning and Memory**, American Psychological Association (APA), v. 6, n. 2, p. 174–215, 1980. Disponível em: <https://doi.org/10.1037/0278-7393.6.2.174>.
- SOKOLOVA, M.; JAPKOWICZ, N.; SZPAKOWICZ, S. Beyond accuracy, f-score and ROC: A family of discriminant measures for performance evaluation. In: **Lecture Notes in Computer Science**. Springer Berlin Heidelberg, 2006. p. 1015–1021. Disponível em: https://doi.org/10.1007/11941439_114.
- SPRUSTON, N. Pyramidal neurons: dendritic structure and synaptic integration. **Nature Reviews Neuroscience**, Springer Science and Business Media LLC, v. 9, n. 3, p. 206–221, mar. 2008. Disponível em: <https://doi.org/10.1038/nrn2286>.
- SUYKENS, J.; VANDEWALLE, J. **Neural Processing Letters**, Springer Science and Business Media LLC, v. 9, n. 3, p. 293–300, 1999. Disponível em: <https://doi.org/10.1023/a:1018628609742>.
- SVENSÉN, M.; BISHOP, C. M. **Pattern recognition and machine learning**. [S. l.]: Springer, 2007.
- TABERNA, G. A.; MARINO, M.; GANZETTI, M.; MANTINI, D. Spatial localization of eeg electrodes using 3d scanning. **Journal of neural engineering**, IOP Publishing, v. 16, n. 2, p. 026020, 2019.

TAMBURRO, G.; FIEDLER, P.; STONE, D.; HAUEISEN, J.; COMANI, S. A new ICA-based fingerprint method for the automatic removal of physiological artifacts from EEG recordings. **PeerJ**, PeerJ, v. 6, p. e4380, fev. 2018. Disponível em: <https://doi.org/10.7717/peerj.4380>.

TATUM, W.; RUBBOLI, G.; KAPLAN, P.; MIRSATARI, S.; RADHAKRISHNAN, K.; GLOSS, D.; CABOCLO, L.; DRISLANE, F.; KOUTROUMANIDIS, M.; SCHOMER, D.; TRENITE, D. K.-N.; COOK, M.; BENICZKY, S. Clinical utility of EEG in diagnosing and monitoring epilepsy in adults. **Clinical Neurophysiology**, Elsevier BV, v. 129, n. 5, p. 1056–1082, maio 2018. Disponível em: <https://doi.org/10.1016/j.clinph.2018.01.019>.

THARWAT, A. Classification assessment methods. **Applied Computing and Informatics**, Emerald, v. 17, n. 1, p. 168–192, jul. 2020. Disponível em: <https://doi.org/10.1016/j.aci.2018.08.003>.

TONONI, G.; CIRELLI, C. Time to be SHY? some comments on sleep and synaptic homeostasis. **Neural Plasticity**, Hindawi Limited, v. 2012, p. 1–12, 2012. Disponível em: <https://doi.org/10.1155/2012/415250>.

TZIMOURTA, K. D.; CHRISTOU, V.; TZALLAS, A. T.; GIANNAKEAS, N.; ASTRAKAS, L. G.; ANGELIDIS, P.; TSALIKAKIS, D.; TSIPOURAS, M. G. Machine learning algorithms and statistical approaches for alzheimer's disease analysis based on resting-state EEG recordings: A systematic review. **International Journal of Neural Systems**, World Scientific Pub Co Pte Lt, v. 31, n. 05, p. 2130002, fev. 2021. Disponível em: <https://doi.org/10.1142/s0129065721300023>.

URIGÜEN, J. A.; GARCIA-ZAPIRAIN, B. Eeg artifact removal—state-of-the-art and guidelines. **Journal of neural engineering**, IOP Publishing, v. 12, n. 3, p. 031001, 2015.

VALENTINE, D. **Terminology & waveforms**. 2020. Disponível em: <https://www.learningeeg.com/terminology-and-waveforms>.

VYAZOVSKIY, V. V.; TOBLER, I. Theta activity in the waking EEG is a marker of sleep propensity in the rat. **Brain Research**, Elsevier BV, v. 1050, n. 1-2, p. 64–71, jul. 2005. Disponível em: <https://doi.org/10.1016/j.brainres.2005.05.022>.

WEBB, A. R. **Statistical pattern recognition**. [S. l.]: John Wiley & Sons, 2003.

WEBSTER, J. G. **Medical instrumentation: application and design**. [S. l.]: John Wiley & Sons, 2009.

YU, M.; LI, G.; JIANG, D.; JIANG, G.; TAO, B.; CHEN, D. Hand medical monitoring system based on machine learning and optimal EMG feature set. **Personal and Ubiquitous Computing**, Springer Science and Business Media LLC, ago. 2019. Disponível em: <https://doi.org/10.1007/s00779-019-01285-2>.

ZAZZARO, G.; CUOMO, S.; MARTONE, A.; MONTAQUILA, R. V.; TORALDO, G.; PAVONE, L. EEG signal analysis for epileptic seizures detection by applying data mining techniques. **Internet of Things**, Elsevier BV, v. 14, p. 100048, jun. 2021. Disponível em: <https://doi.org/10.1016/j.iot.2019.03.002>.

ZHANG, X. L.; BEGLEITER, H.; PORJESZ, B.; WANG, W.; LITKE, A. Event related potentials during object recognition tasks. **Brain research bulletin**, Elsevier, v. 38, n. 6, p. 531–538, 1995.

APÊNDICE A – RESULTADOS MÉDIOS CASO CHB01

Tabela 5 – Melhores resultados para o caso CHB01.

		Segmentação 01			
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	94,96%	95,29%	77,88%	88,99%
	LPC	95,31%	95,54%	79,98%	89,97%
	Covariância (intra)	93,43%	93,74%	68,87%	84,49%
	KFE RBF (intra) - P	90,99%	90,85%	96,43%	92,72%
	RENYI RBF (intra)	87,71%	87,48%	96,43%	90,45%
CLG	Covariância (inter)	92,87%	92,99%	85,47%	90,37%
	LPC	95,42%	95,63%	80,68%	90,29%
	Covariância (intra)	92,29%	92,70%	63,62%	81,65%
	KFE RBF (intra) - P	90,96%	90,80%	94,56%	92,09%
	RENYI RBF (intra)	89,32%	89,08%	97,50%	91,88%
		Segmentação 02			
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	96,96%	97,50%	64,45%	84,78%
	LPC	91,98%	91,91%	90,26%	91,38%
	Covariância (intra)	92,74%	93,15%	64,80%	82,42%
	RENYI RBF (inter)	82,42%	82,12%	97,37%	87,02%
	KFE RBF (intra) - P	93,21%	93,27%	90,22%	92,22%
RENYI RBF (intra)	93,20%	93,23%	91,73%	92,72%	
CLG	Covariância (inter)	97,29%	97,72%	75,31%	89,46%
	LPC	91,42%	91,40%	84,31%	88,98%
	Covariância (intra)	95,19%	95,59%	70,92%	86,41%
	KFE RBF (inter) - P	94,21%	94,37%	86,62%	91,66%
	RENYI RBF (inter)	91,96%	91,96%	92,10%	92,01%
	KFE RBF (intra) - P	93,73%	93,76%	90,23%	92,56%
RENYI RBF (intra)	93,08%	93,04%	95,41%	93,84%	
		Segmentação 03			
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	97,03%	97,37%	81,83%	91,78%
	LPC	96,14%	96,18%	90,12%	94,10%
	Covariância (intra)	90,21%	90,44%	64,38%	80,69%
	KFE RBF (intra) - P	92,10%	92,12%	91,10%	91,77%
	RENYI RBF (intra)	89,33%	89,21%	95,22%	91,21%
CLG	Covariância (inter)	96,67%	96,99%	81,79%	91,53%
	LPC	94,24%	94,12%	96,52%	94,96%
	Covariância (intra)	93,87%	94,24%	70,62%	85,48%
	KFE RBF (inter) - P	93,62%	93,75%	85,51%	90,88%
	RENYI RBF (inter)	90,43%	90,39%	91,65%	90,82%
	KFE RBF (intra) - P	90,58%	90,57%	90,20%	90,45%
RENYI RBF (intra)	90,53%	90,43%	96,69%	92,50%	
		Segmentação 04			
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	96,20%	96,20%	93,82%	95,40%
	LPC	98,51%	98,49%	99,70%	98,90%
	Covariância (intra)	97,78%	97,94%	85,78%	93,66%
	KFE RBF (intra) - P	92,65%	92,57%	95,96%	93,71%
	RENYI RBF (intra)	94,06%	94,08%	90,10%	92,73%
CLG	Covariância (inter)	96,20%	96,57%	76,97%	89,42%
	LPC	98,11%	98,19%	83,27%	92,91%
	Covariância (intra)	94,66%	94,38%	71,35%	86,07%
	KFE RBF (inter) - P	92,19%	93,65%	84,92%	90,17%
	RENYI RBF (inter)	89,30%	90,24%	94,85%	91,43%
	RENYI'S POLY (inter)	96,13%	96,86%	76,20%	89,19%
	KFE RBF (intra) - P	91,05%	91,97%	94,23%	92,41%
	KFE POLY (intra) - P	94,75%	96,21%	72,83%	87,24%
	RENYI RBF (intra)	90,19%	90,67%	96,96%	92,55%
RENYI POLY (intra)	95,78%	97,73%	76,41%	89,43%	

Fonte: Elaborado pelo Autor.

APÊNDICE B – RESULTADOS MÉDIOS CASO CHB03

Tabela 6 – Melhores resultados para o caso **CHB03**.

Segmentação 01					
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	97,38%	97,76%	73,64%	88,83%
	LPC	94,87%	94,88%	94,12%	94,62%
	Covariância (intra)	88,94%	89,45%	57,71%	77,15%
	KFE RBF (intra) - P	93,69%	93,78%	87,48%	91,60%
	RENYI RBF (intra)	93,64%	93,62%	94,38%	93,88%
CLG	Covariância (inter)	96,52%	96,84%	76,98%	89,61%
	LPC	95,11%	95,11%	95,31%	95,18%
	Covariância (intra)	93,83%	94,19%	71,13%	85,67%
	KFE RBF (intra) - P	91,73%	91,65%	96,46%	93,25%
	RENYI RBF (intra)	93,64%	93,62%	94,38%	93,88%
Segmentação 02					
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	96,65%	96,85%	84,21%	92,38%
	LPC	96,83%	96,94%	90,45%	94,69%
	Covariância (intra)	93,43%	94,09%	52,32%	77,19%
	KFE RBF (intra) - P	96,41%	96,83%	69,86%	86,72%
	RENYI RBF (intra)	97,40%	97,53%	88,64%	94,43%
CLG	Covariância (inter)	96,32%	96,54%	81,49%	91,17%
	LPC	93,59%	93,57%	94,18%	93,78%
	Covariância (intra)	96,25%	96,62%	73,16%	87,95%
	RENYI RBF (inter)	86,24%	86,42%	74,66%	82,25%
	KFE RBF (intra) - P	93,37%	93,52%	82,67%	89,71%
RENYI RBF (intra)	93,43%	93,51%	87,31%	91,37%	
Segmentação 03					
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	96,83%	97,16%	75,99%	89,42%
	LPC	97,75%	97,81%	95,00%	96,84%
	Covariância (intra)	91,05%	91,72%	48,50%	73,99%
	KFE RBF (intra) - P	95,41%	95,62%	82,15%	90,83%
	RENYI RBF (intra)	95,10%	95,14%	92,31%	94,18%
CLG	Covariância (inter)	97,33%	97,59%	80,56%	91,46%
	LPC	96,96%	96,98%	96,30%	96,75%
	Covariância (intra)	94,28%	94,74%	64,93%	83,39%
	RENYI RBF (inter)	84,44%	84,48%	81,13%	83,33%
	KFE RBF (intra) - P	91,77%	91,83%	87,14%	90,22%
RENYI RBF (intra)	93,61%	93,55%	96,91%	94,68%	
Segmentação 04					
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	97,90%	98,24%	76,86%	90,42%
	LPC	97,54%	97,61%	93,34%	96,14%
	Covariância (intra)	93,07%	93,74%	51,91%	76,80%
	KFE RBF (intra) - P	95,16%	95,34%	83,65%	91,21%
	RENYI RBF (intra)	96,33%	96,42%	89,98%	94,19%
CLG	Covariância (inter)	98,26%	98,57%	78,77%	91,38%
	LPC	95,23%	95,21%	96,03%	95,49%
	Covariância (intra)	95,71%	96,09%	71,92%	87,13%
	RENYI RBF (intra)	92,97%	92,96%	92,83%	92,92%
	RENYI POLY (intra)	95,60%	95,62%	93,83%	95,01%

Fonte: Elaborado pelo Autor.

APÊNDICE C – RESULTADOS MÉDIOS CASO CHB14

Tabela 7 – Melhores resultados para o caso **CHB14**.

Segmentação 01					
	Acurácia	Especificidade	Sensibilidade	M.G	
OLS	Covariância (inter)	71,60%	71,65%	61,40%	68,04%
	LPC	92,71%	92,75%	86,46%	90,59%
	Covariância (intra)	83,70%	83,86%	57,81%	74,03%
	KFE RBF (intra) - P	92,98%	93,20%	61,68%	81,16%
	RENYI RBF (intra)	95,62%	95,81%	71,12%	86,69%
CLG	Covariância (inter)	69,65%	69,65%	64,55%	67,91%
	LPC	91,62%	91,68%	80,43%	87,75%
	Covariância (intra)	76,93%	77,08%	52,35%	67,71%
	KFE RBF (intra) - P	79,67%	79,74%	62,88%	73,65%
	RENYI RBF (intra)	89,84%	89,96%	71,61%	83,34%
Segmentação 02					
	Acurácia	Especificidade	Sensibilidade	M.G	
OLS	Covariância (inter)	59,37%	59,46%	42,70%	53,22%
	LPC	42,78%	42,43%	85,60%	53,76%
	Covariância (intra)	74,85%	74,99%	53,90%	67,13%
	KFE RBF (intra) - P	92,55%	92,71%	73,04%	85,58%
	RENYI RBF (intra)	95,29%	95,51%	68,85%	85,57%
CLG	Covariância (inter)	57,40%	57,42%	45,10%	52,97%
	LPC	46,12%	45,88%	75,07%	54,16%
	Covariância (intra)	70,06%	70,24%	48,53%	62,04%
	KFE RBF (intra) - P	79,07%	79,15%	69,66%	75,82%
	RENYI RBF (intra)	89,67%	89,78%	74,20%	84,22%
Segmentação 03					
	Acurácia	Especificidade	Sensibilidade	M.G	
OLS	Covariância (inter)	61,19%	61,15%	68,57%	63,54%
	LPC	55,59%	55,32%	89,51%	65,05%
	Covariância (intra)	78,95%	79,02%	70,95%	76,21%
	KFE RBF (intra) - P	88,13%	88,20%	79,50%	85,17%
	RENYI RBF (intra)	92,64%	92,76%	77,57%	87,36%
CLG	Covariância (inter)	64,48%	64,38%	81,48%	69,68%
	LPC	67,08%	67,03%	72,44%	68,80%
	Covariância (intra)	74,02%	74,16%	57,11%	67,93%
	KFE RBF (intra) - P	67,86%	67,86%	68,11%	67,94%
	RENYI RBF (intra)	84,82%	84,83%	85,45%	85,03%
Segmentação 04					
	Acurácia	Especificidade	Sensibilidade	M.G	
OLS	Covariância (inter)	73,45%	73,55%	56,30%	67,25%
	LPC	94,10%	94,14%	86,66%	91,56%
	Covariância (intra)	84,07%	84,25%	52,37%	71,85%
	KFE RBF (intra) - P	92,07%	92,21%	74,37%	85,79%
	RENYI RBF (intra)	94,81%	94,96%	74,20%	87,42%
CLG	Covariância (inter)	72,56%	72,63%	60,02%	68,13%
	LPC	93,92%	94,08%	70,37%	85,35%
	Covariância (intra)	76,67%	76,77%	59,81%	70,61%
	KFE RBF (intra) - P	71,22%	71,25%	65,70%	69,34%
	RENYI POLY (intra)	88,64%	88,75%	74,34%	83,63%

Fonte: Elaborado pelo Autor.

APÊNDICE D – RESULTADOS MÉDIOS CASO CHB18

Tabela 8 – Melhores resultados para o caso **CHB18**.

Segmentação 01					
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	92,69%	93,52%	44,49%	72,79%
	LPC	52,01%	51,50%	76,70%	59,00%
	Covariância (intra)	62,59%	62,58%	56,22%	60,39%
	KFE RBF (intra) - P	85,64%	85,96%	61,82%	76,92%
	RENYI RBF (intra)	86,89%	87,00%	72,84%	81,97%
CLG	Covariância (inter)	84,66%	85,24%	54,92%	73,45%
	LPC	60,60%	60,22%	78,43%	65,90%
	Covariância (intra)	71,30%	71,34%	61,62%	67,93%
	KFE RBF (intra) - P	80,09%	80,22%	69,09%	76,28%
	RENYI RBF (intra)	75,29%	75,18%	75,13%	75,20%
Segmentação 02					
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	67,94%	68,24%	47,54%	60,41%
	LPC	67,11%	67,49%	45,08%	58,89%
	Covariância (intra)	33,07%	32,41%	80,06%	44,11%
	KFE RBF (intra) - P	86,82%	87,22%	52,82%	73,68%
	RENYI RBF (intra)	87,83%	88,22%	53,84%	74,72%
CLG	Covariância (inter)	80,79%	81,30%	44,17%	66,20%
	LPC	66,83%	67,01%	61,58%	65,09%
	Covariância (intra)	72,99%	73,44%	33,84%	56,61%
	KFE RBF (intra) - P	82,65%	82,83%	61,08%	74,78%
	RENYI RBF (intra)	83,07%	83,19%	66,13%	77,03%
Segmentação 03					
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	93,75%	94,50%	47,23%	74,79%
	LPC	75,99%	76,19%	55,97%	68,68%
	Covariância (intra)	70,00%	70,39%	42,35%	59,31%
	KFE RBF (intra) - P	89,54%	90,09%	55,53%	76,51%
	RENYI RBF (intra)	91,98%	92,76%	46,11%	73,27%
CLG	Covariância (inter)	93,28%	93,96%	51,49%	76,70%
	LPC	69,75%	69,66%	67,77%	69,06%
	Covariância (intra)	68,29%	68,37%	57,24%	64,41%
	KFE RBF (intra) - P	87,10%	87,52%	61,91%	77,86%
	RENYI RBF (intra)	88,27%	88,46%	74,08%	83,32%
Segmentação 04					
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	68,11%	68,59%	42,63%	58,40%
	LPC	71,65%	71,98%	46,10%	61,95%
	Covariância (intra)	59,93%	59,77%	54,93%	58,16%
	KFE RBF (intra) - P	83,43%	83,94%	43,67%	67,37%
	RENYI RBF (intra)	85,94%	86,29%	55,66%	74,46%
CLG	Covariância (inter)	77,83%	78,37%	40,55%	62,77%
	LPC	80,38%	80,80%	46,14%	66,92%
	Covariância (intra)	60,94%	60,78%	56,56%	59,39%
	KFE RBF (intra) - P	83,14%	83,39%	59,21%	74,32%
	RENYI POLY (intra)	80,44%	80,40%	72,49%	77,69%

Fonte: Elaborado pelo Autor.

APÊNDICE E – RESULTADOS MÉDIOS CASO CHB24

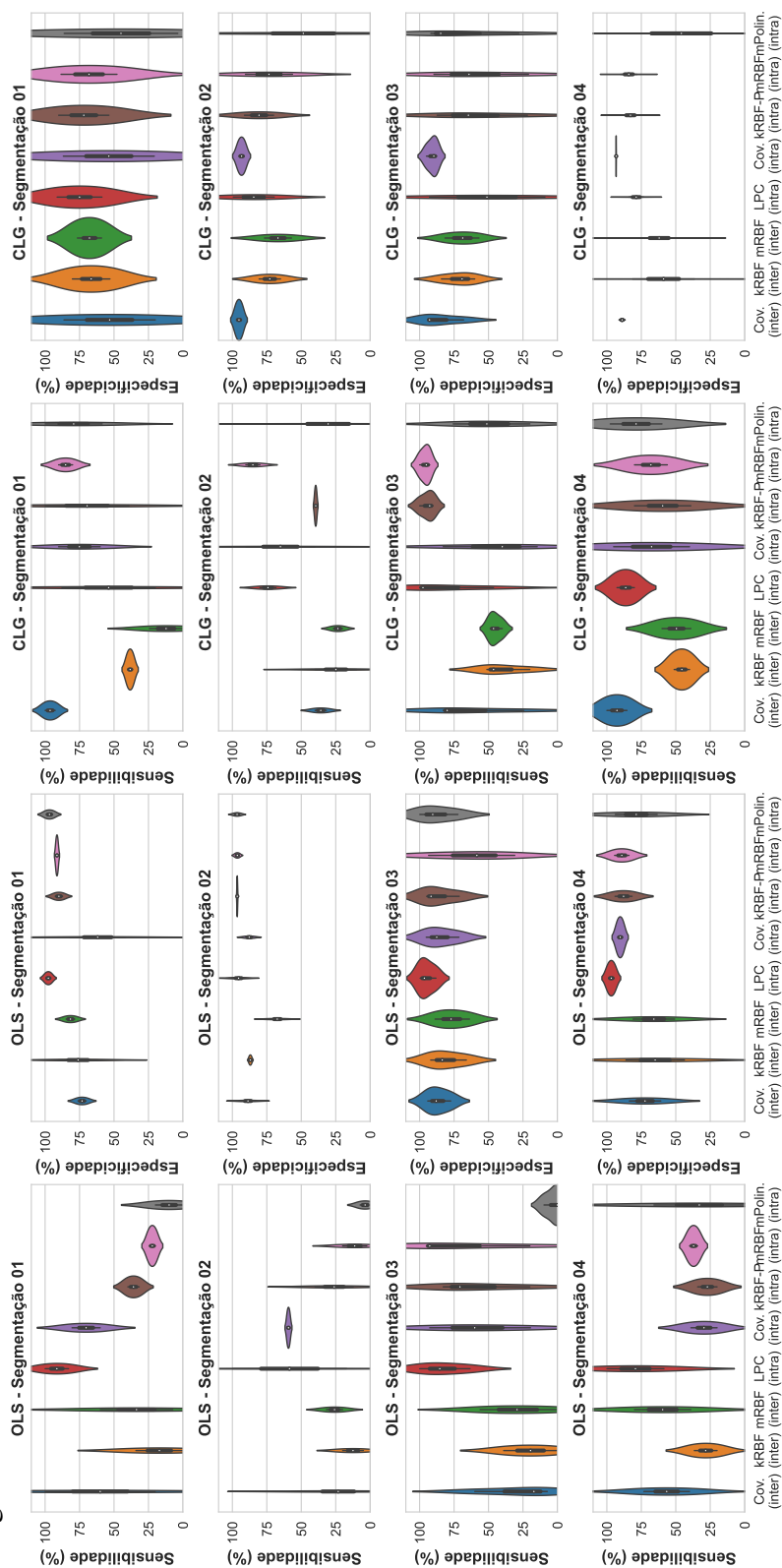
Tabela 9 – Melhores resultados para o caso **CHB24**.

		Segmentação 01			
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	93,40%	93,99%	48,36%	75,16%
	LPC	96,88%	97,27%	65,02%	84,94%
	Covariância (intra)	88,64%	89,16%	45,68%	71,20%
	KFE RBF (intra) - P	91,99%	92,27%	70,87%	84,42%
	RENYI RBF (intra)	94,27%	94,56%	71,62%	86,11%
CLG	Covariância (inter)	91,70%	92,15%	58,07%	78,87%
	LPC	91,86%	92,20%	65,06%	81,98%
	Covariância (intra)	90,21%	90,69%	54,09%	76,20%
	KFE RBF (intra) - P	94,02%	94,40%	67,22%	84,18%
	RENYI RBF (intra)	93,28%	93,64%	66,55%	83,46%
		Segmentação 02			
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	63,90%	63,87%	54,20%	60,48%
	LPC	48,85%	48,35%	88,06%	59,25%
	Covariância (intra)	74,33%	74,31%	59,53%	69,02%
	KFE RBF (intra) - P	91,58%	91,71%	75,80%	86,03%
	RENYI RBF (intra)	93,03%	93,19%	75,49%	86,82%
CLG	Covariância (inter)	67,81%	77,34%	55,50%	66,27%
	LPC	51,48%	51,13%	83,39%	60,32%
	Covariância (intra)	70,15%	70,20%	45,72%	60,84%
	KFE RBF (intra) - P	90,74%	90,83%	77,89%	86,26%
	RENYI RBF (intra)	91,04%	91,12%	78,00%	86,49%
		Segmentação 03			
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	86,45%	86,93%	49,06%	71,70%
	LPC	58,32%	58,13%	80,86%	64,96%
	Covariância (intra)	81,57%	81,82%	53,23%	70,82%
	KFE RBF (intra) - P	87,71%	87,89%	71,04%	81,81%
	RENYI RBF (intra)	90,38%	90,57%	72,55%	84,06%
CLG	Covariância (inter)	84,84%	85,26%	50,94%	71,69%
	LPC	52,01%	51,71%	83,65%	60,82%
	Covariância (intra)	84,49%	84,88%	53,34%	72,59%
	KFE RBF (intra) - P	87,78%	87,89%	75,48%	83,51%
	RENYI RBF (intra)	89,21%	89,36%	74,69%	84,13%
		Segmentação 04			
		Acurácia	Especificidade	Sensibilidade	M.G
OLS	Covariância (inter)	90,34%	90,81%	48,05%	73,32%
	LPC	97,35%	97,69%	64,30%	84,88%
	Covariância (intra)	82,13%	82,35%	51,90%	70,54%
	KFE RBF (intra) - P	95,06%	95,26%	74,76%	87,81%
	RENYI RBF (intra)	94,88%	95,09%	73,35%	87,14%
CLG	Covariância (inter)	87,03%	87,44%	47,58%	71,28%
	LPC	87,56%	87,62%	75,49%	83,35%
	Covariância (intra)	86,35%	86,69%	46,65%	70,42%
	KFE RBF (intra) - P	93,54%	93,72%	73,70%	86,45%
	RENYI POLY (intra)	93,24%	93,37%	75,72%	87,03%

Fonte: Elaborado pelo Autor.

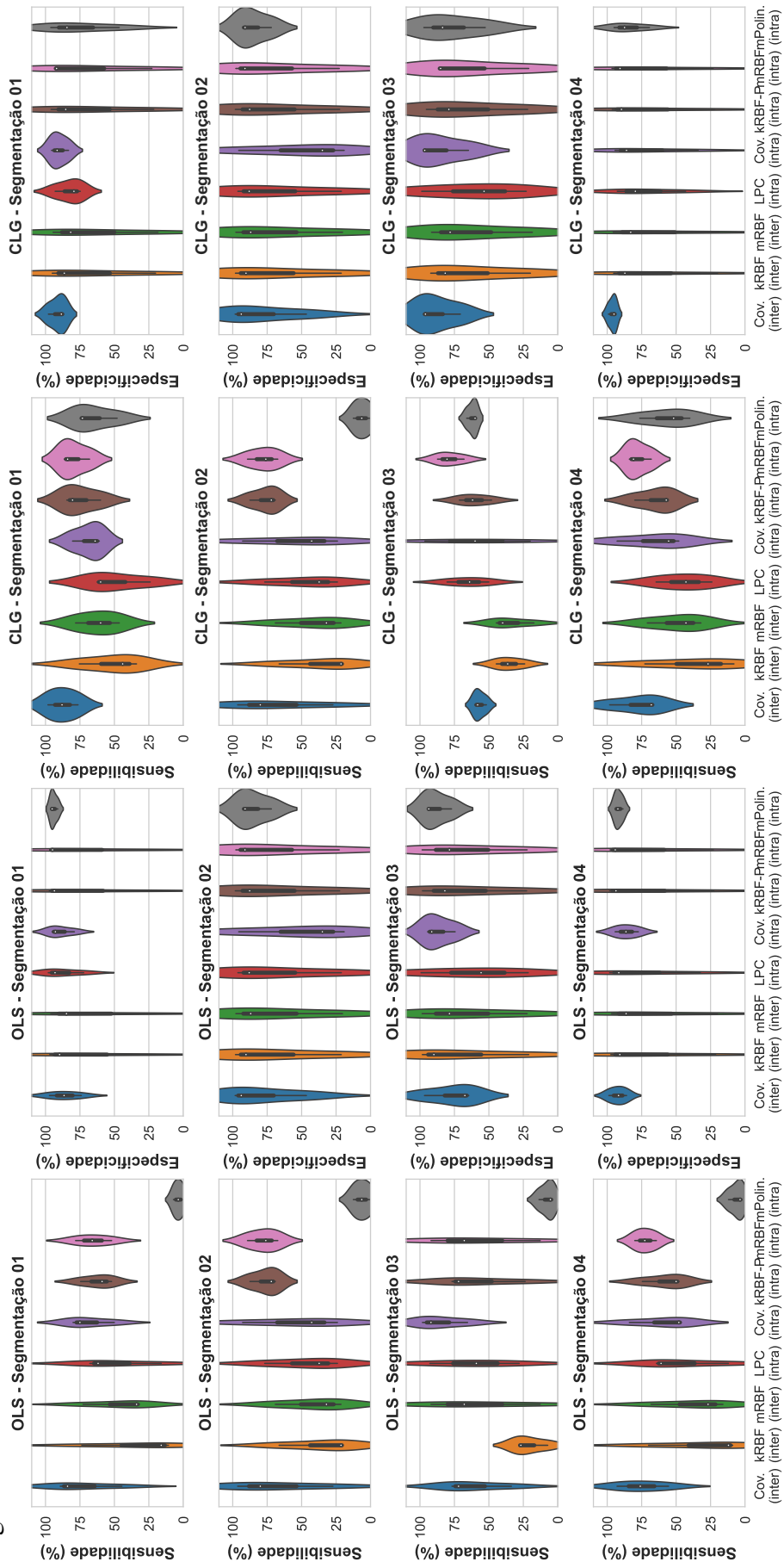
APÊNDICE F – DE MAIS RESULTADOS

Figura 86 – Resultados do caso CHB02.



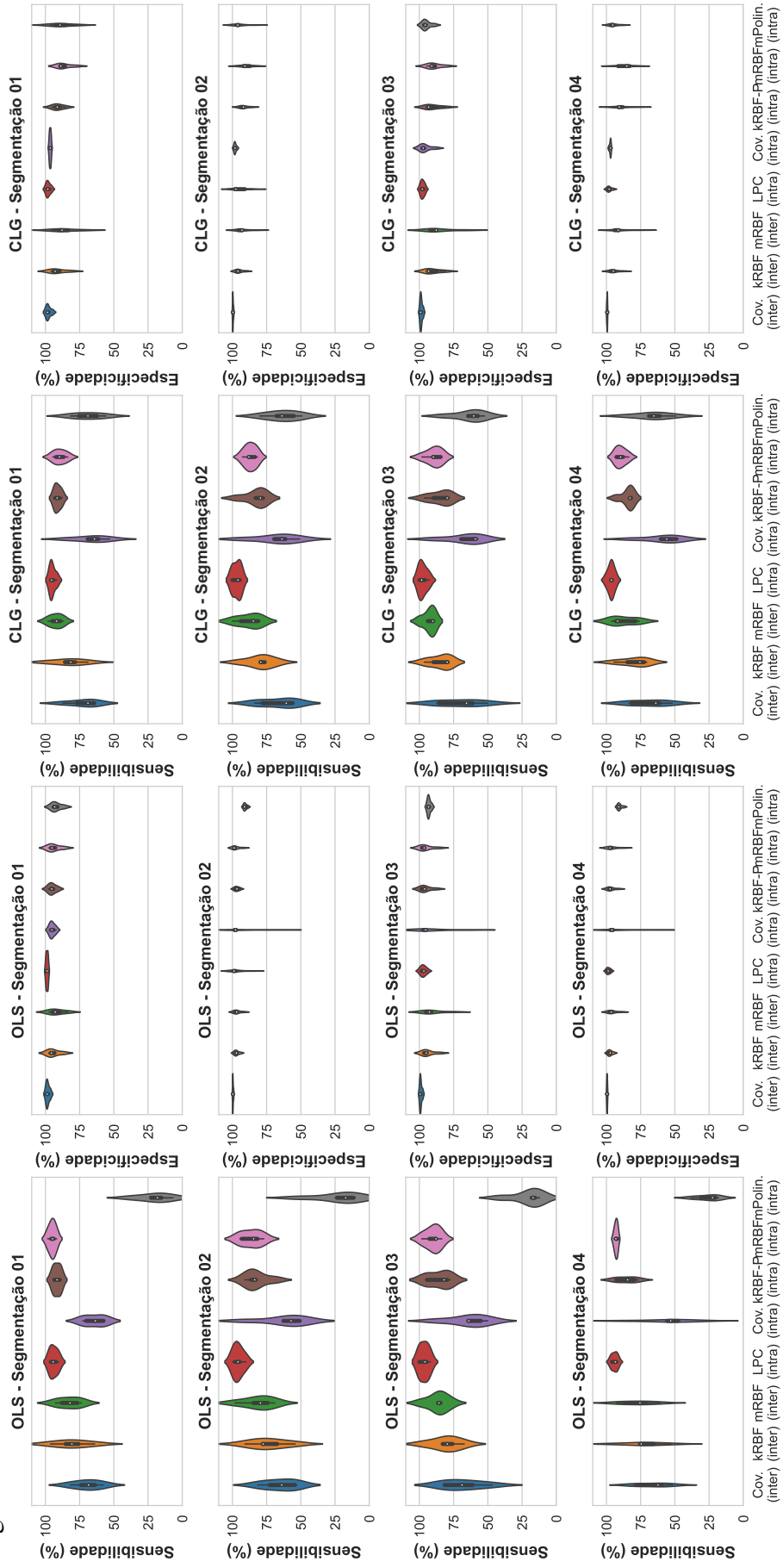
Fonte: Elaborado pelo Autor.

Figura 87 – Resultados do caso CHB04.



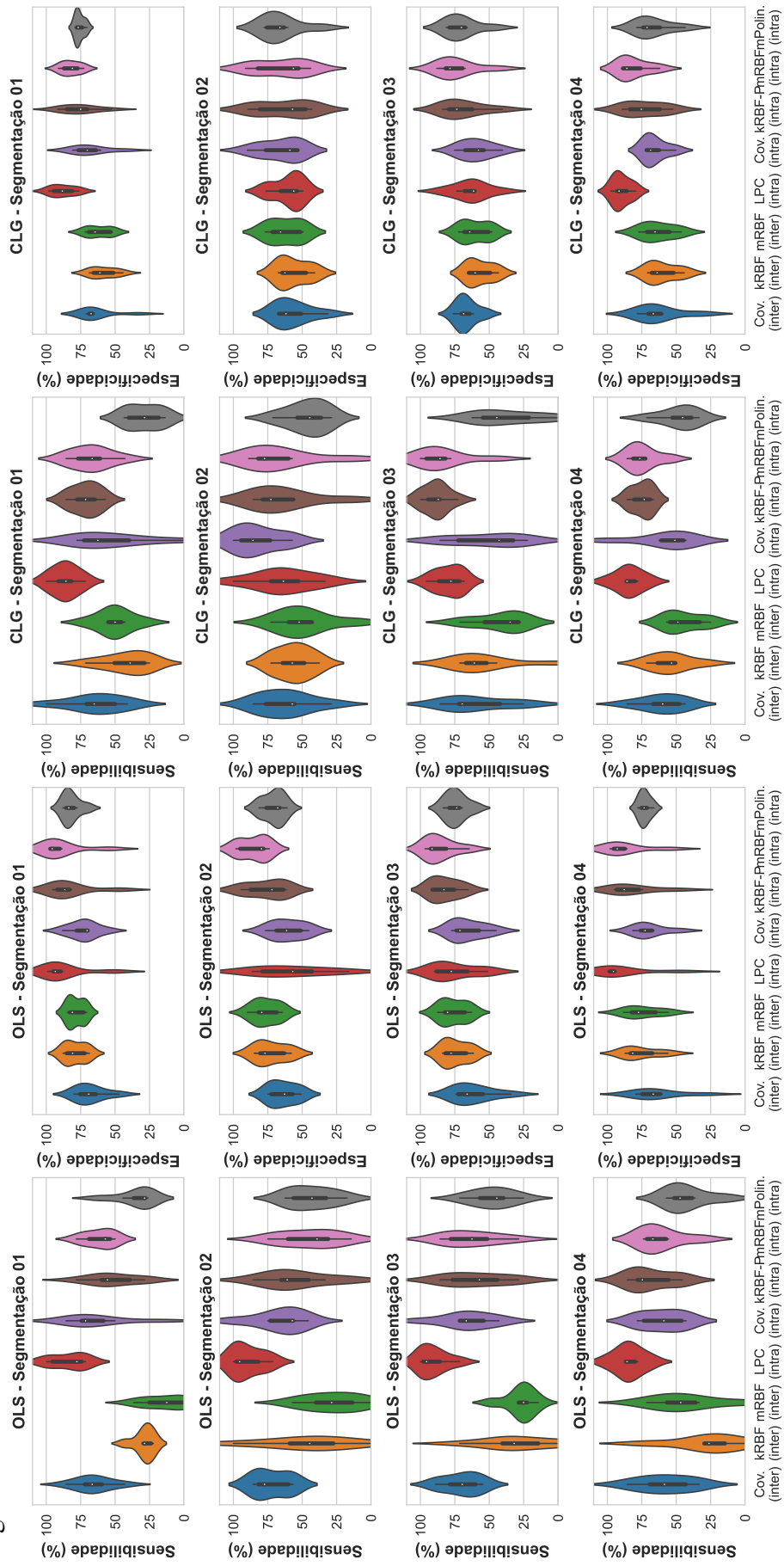
Fonte: Elaborado pelo Autor.

Figura 88 – Resultados do caso CHB05.



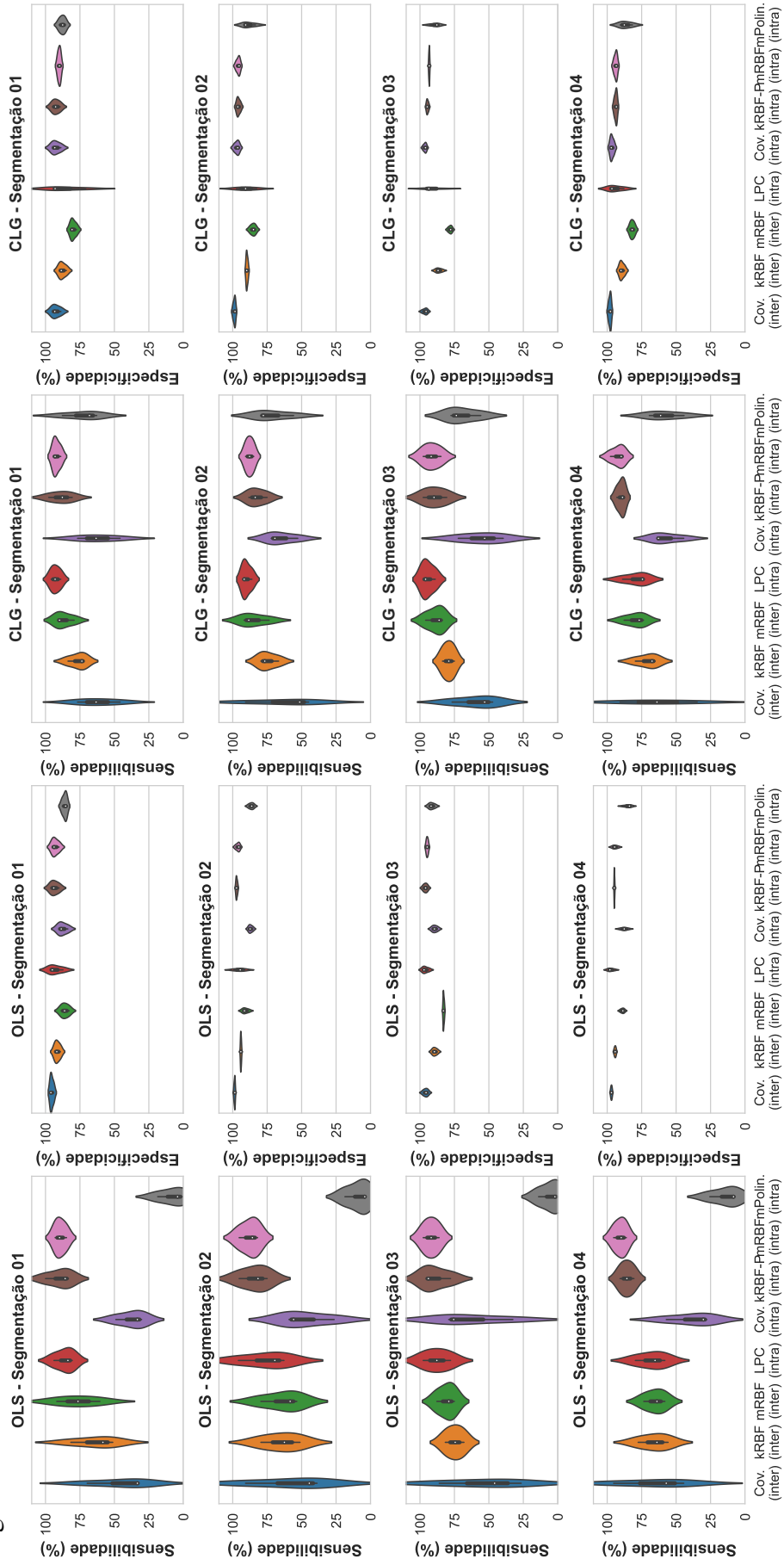
Fonte: Elaborado pelo Autor.

Figura 89 – Resultados do caso CHB06.



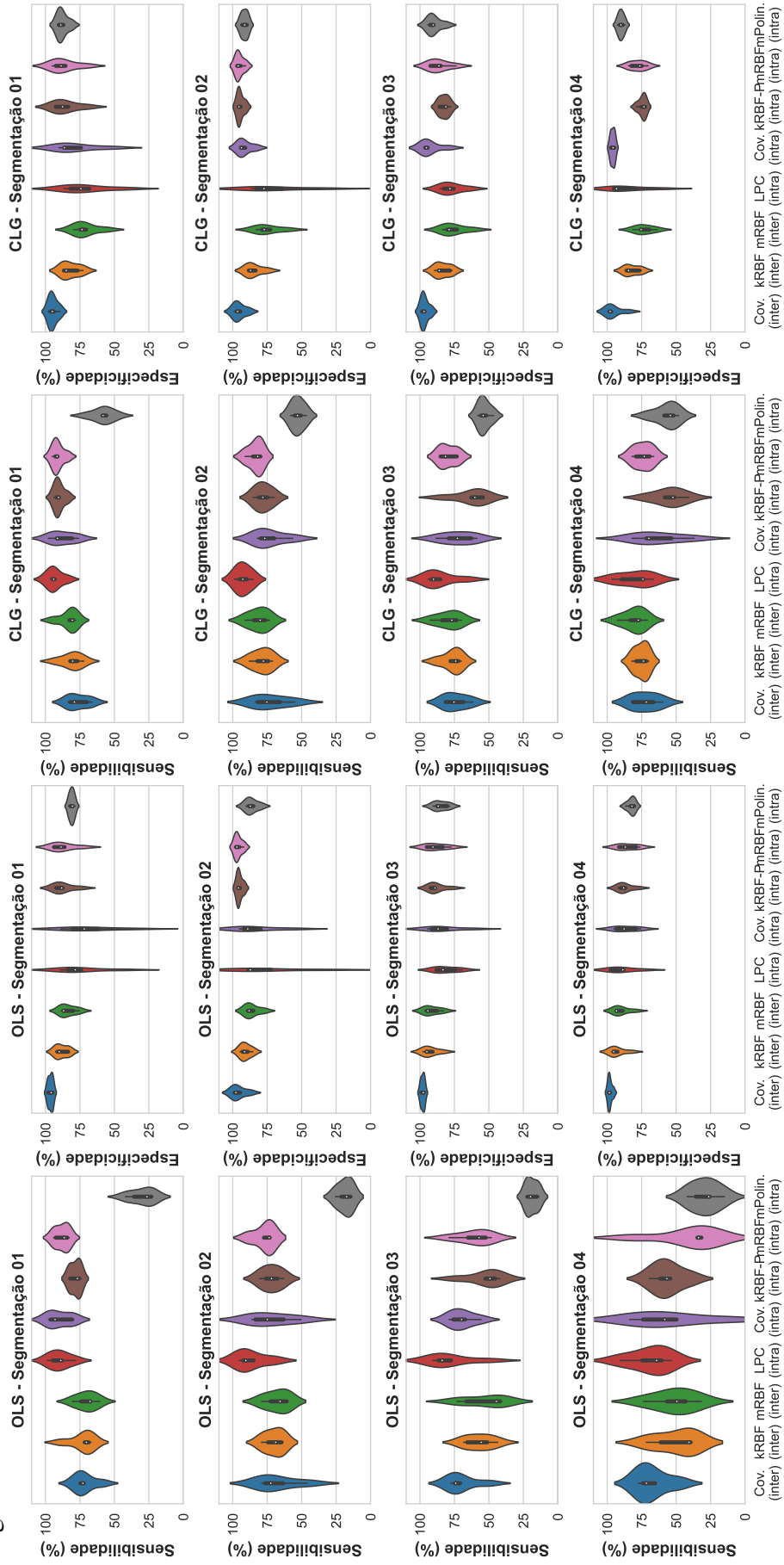
Fonte: Elaborado pelo Autor.

Figura 90 – Resultados do caso CHB07.



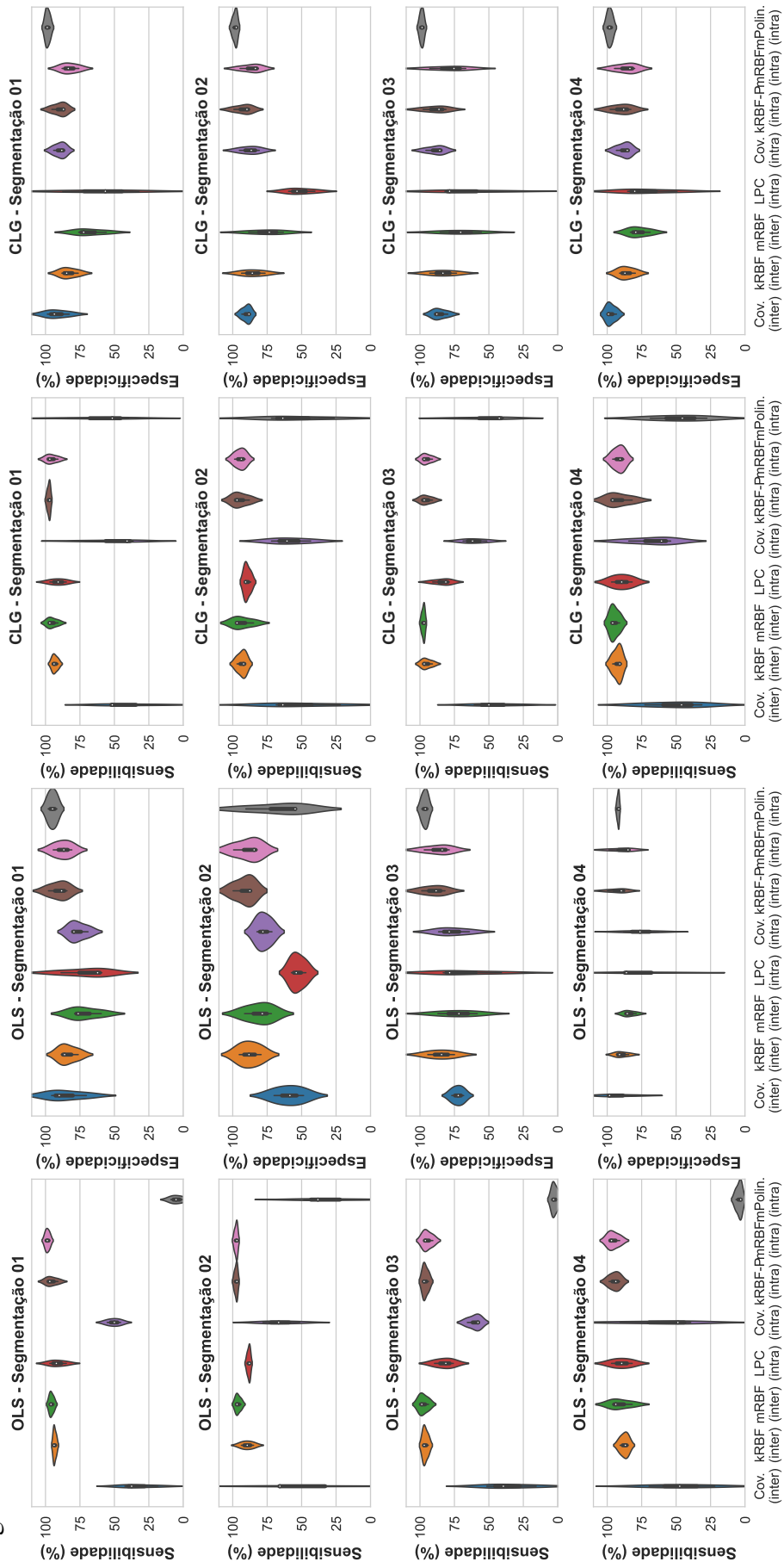
Fonte: Elaborado pelo Autor.

Figura 91 – Resultados do caso CHB08.



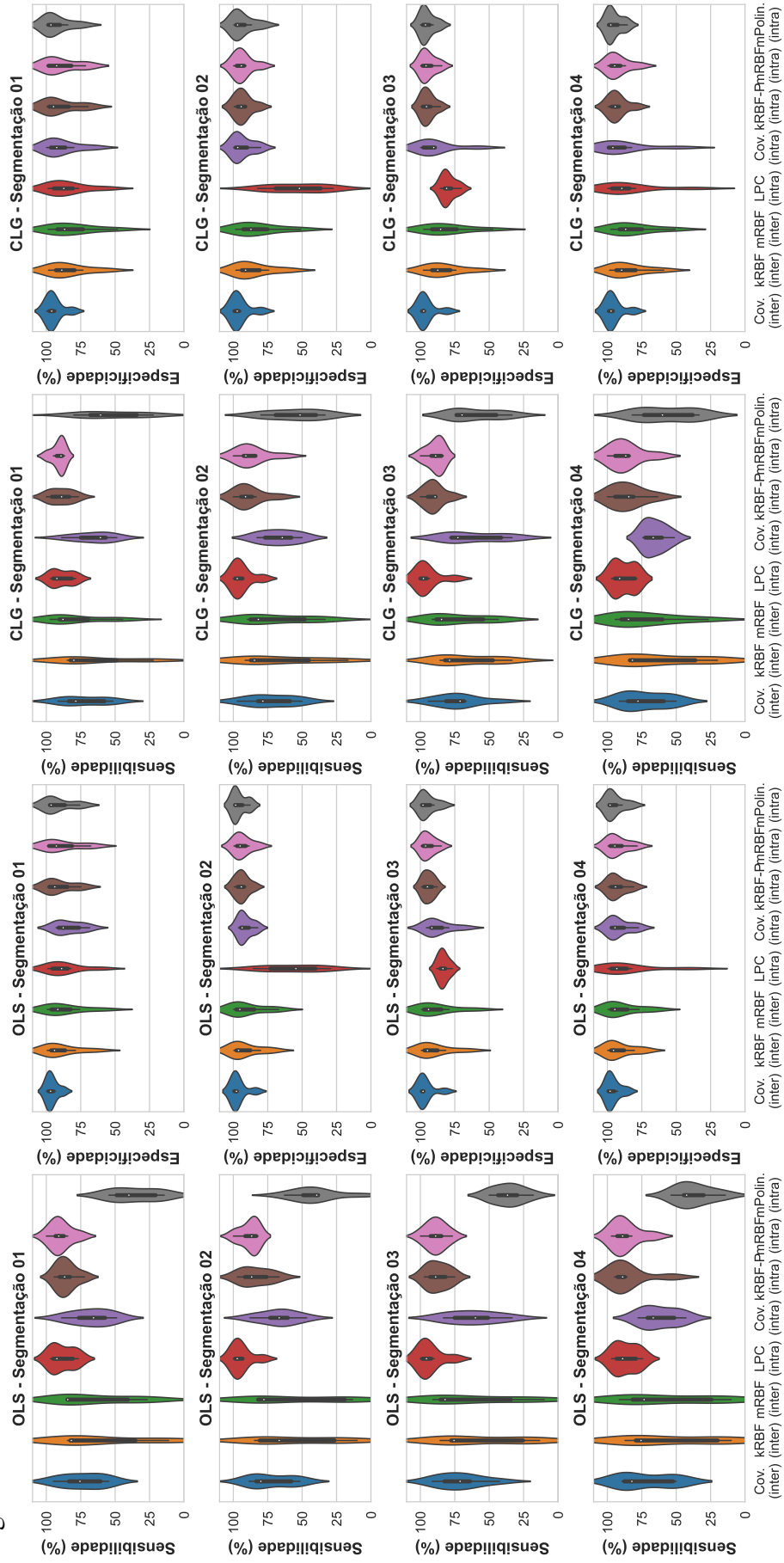
Fonte: Elaborado pelo Autor.

Figura 92 – Resultados do caso CHB09.



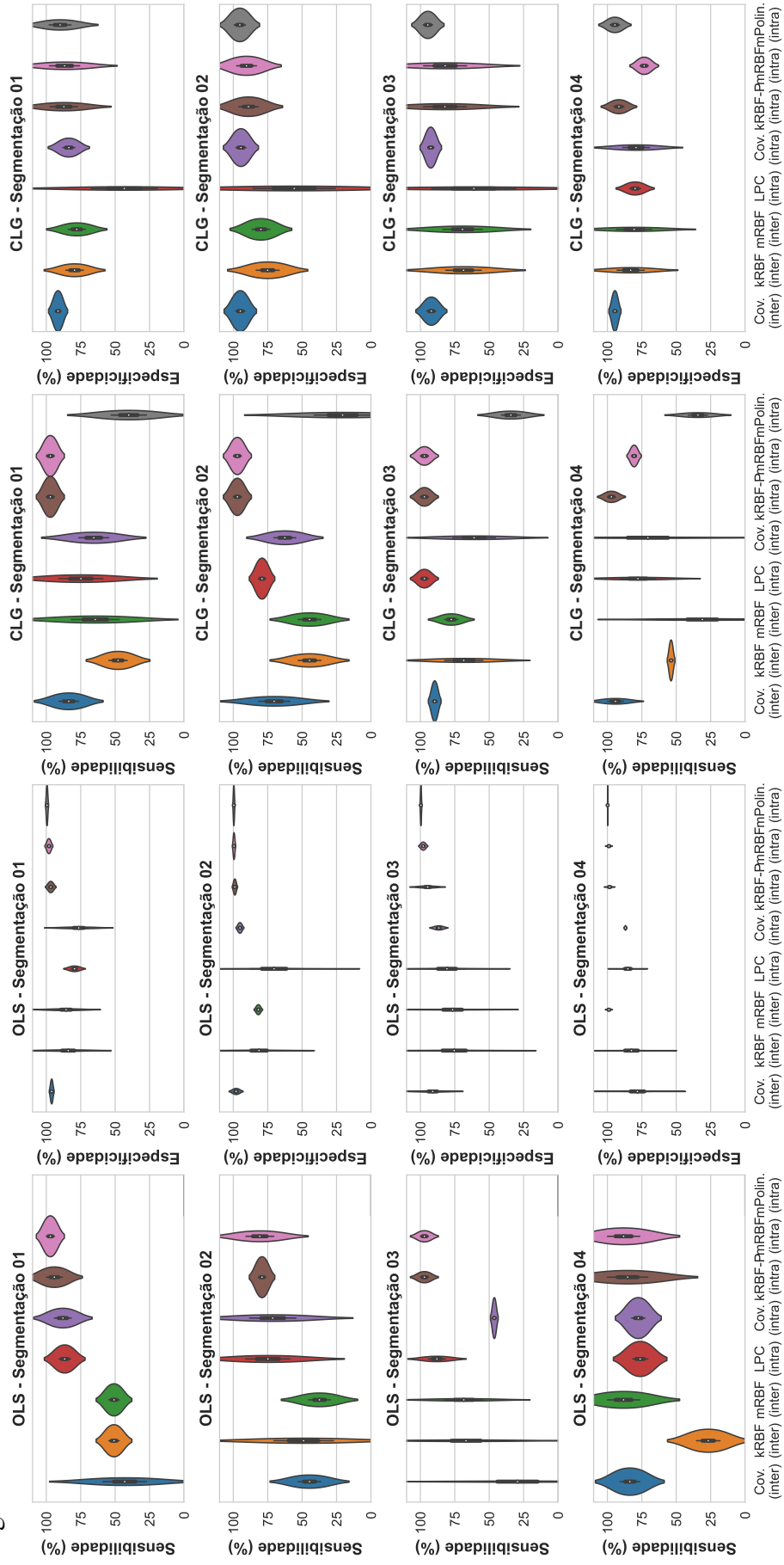
Fonte: Elaborado pelo Autor.

Figura 93 – Resultados do caso CHB10.



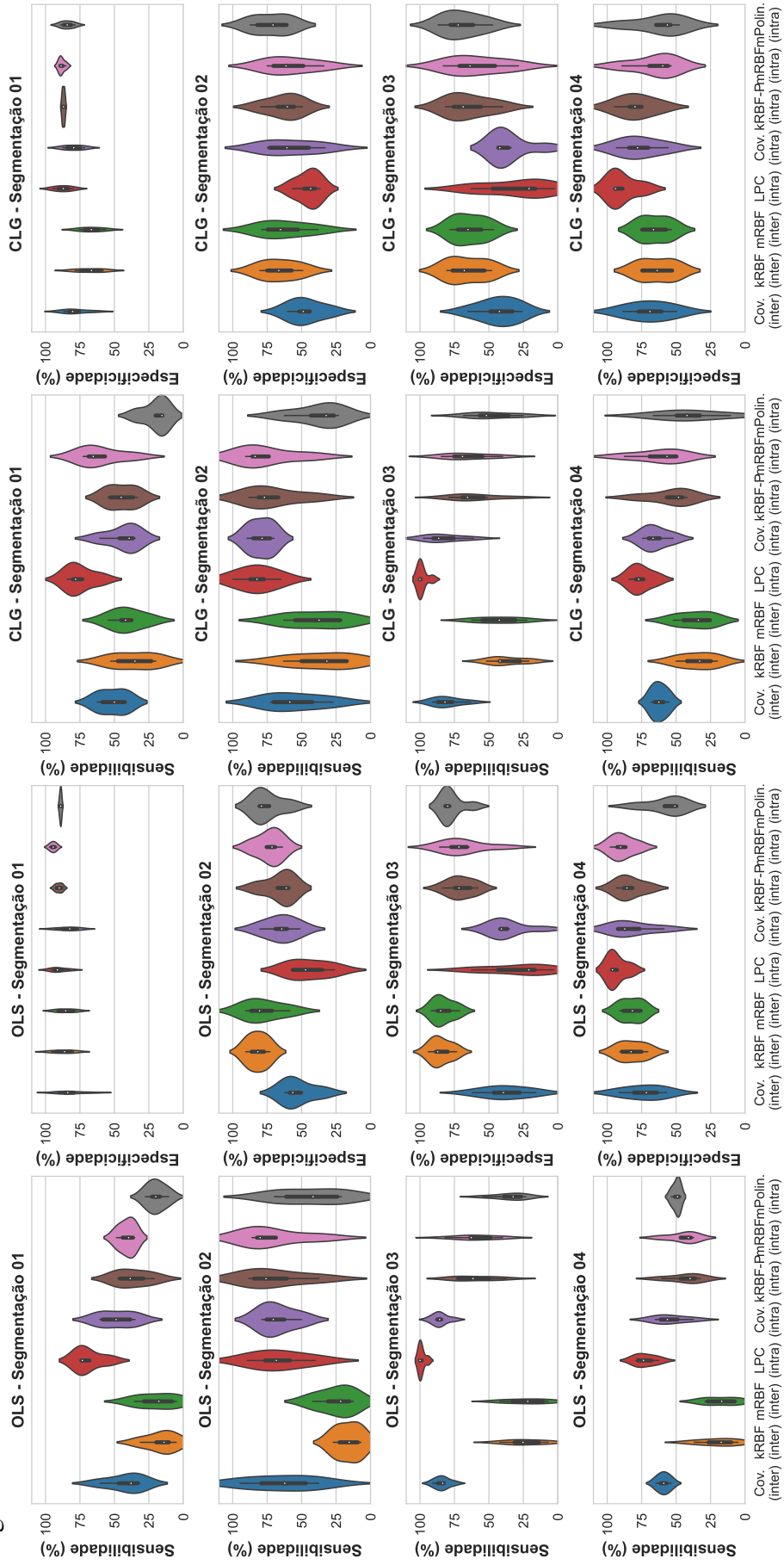
Fonte: Elaborado pelo Autor.

Figura 94 – Resultados do caso CHB11.



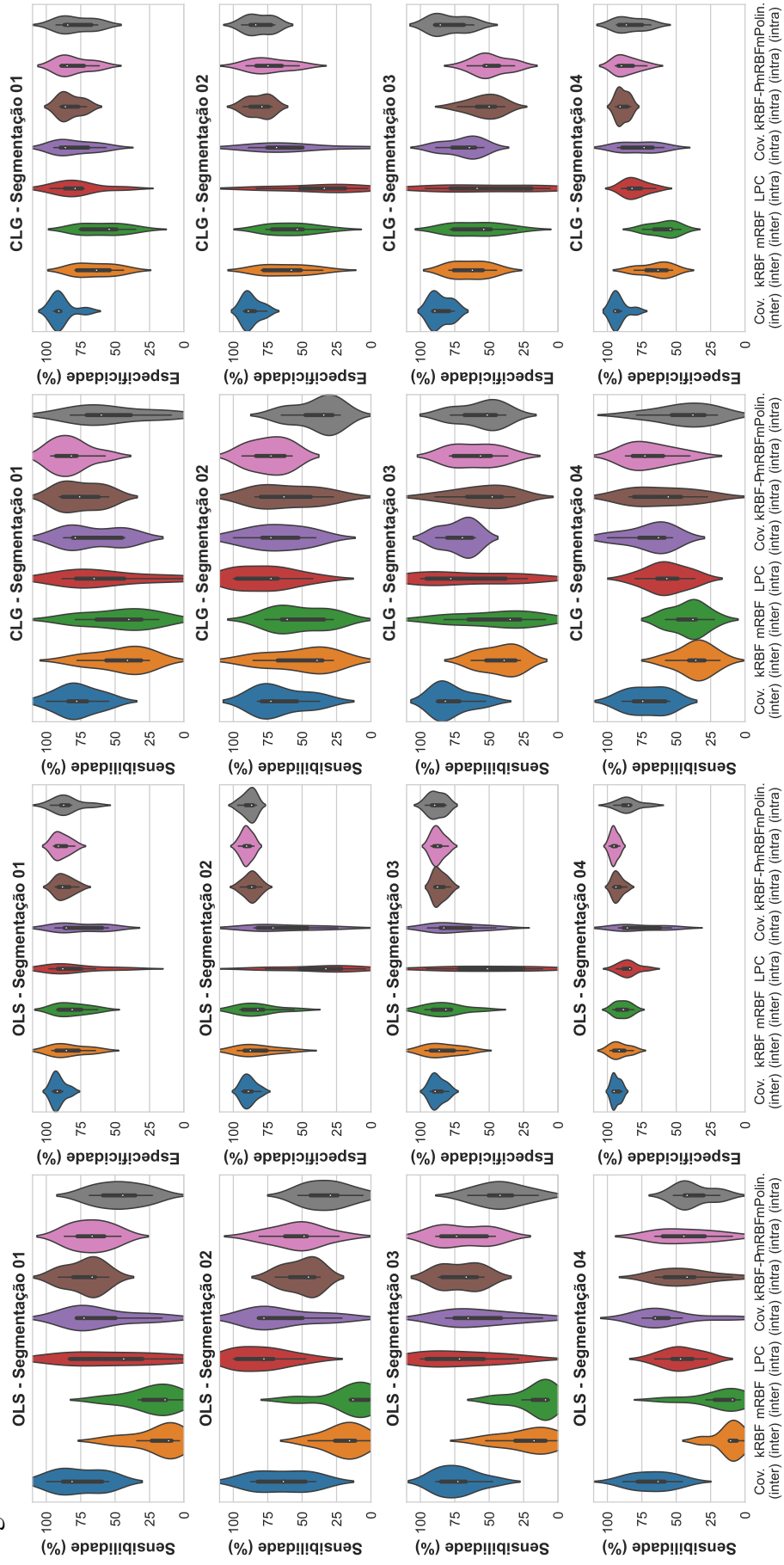
Fonte: Elaborado pelo Autor.

Figura 95 – Resultados do caso CHB12.



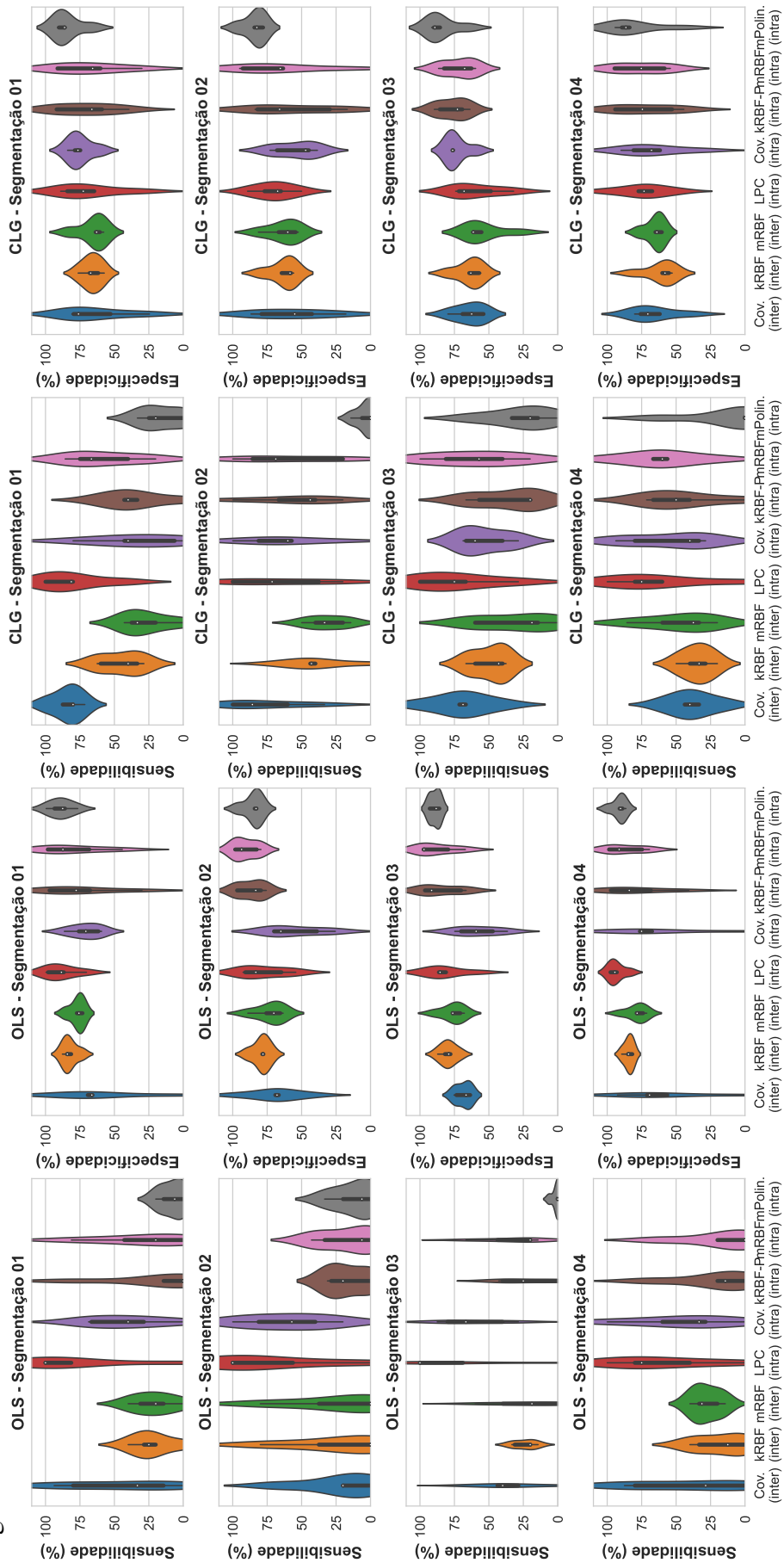
Fonte: Elaborado pelo Autor.

Figura 96 – Resultados do caso CHB13.



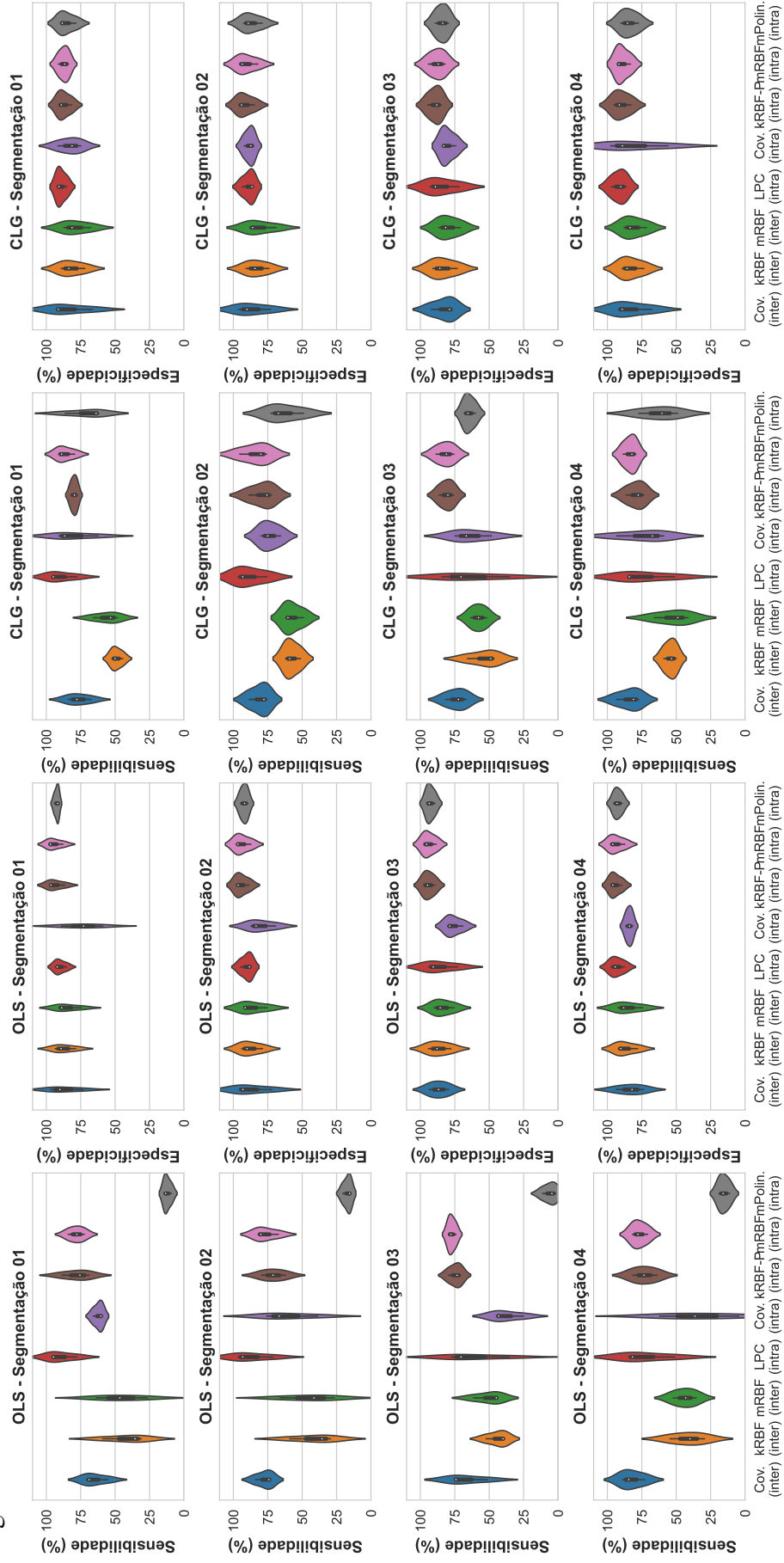
Fonte: Elaborado pelo Autor.

Figura 97 – Resultados do caso CHB16.



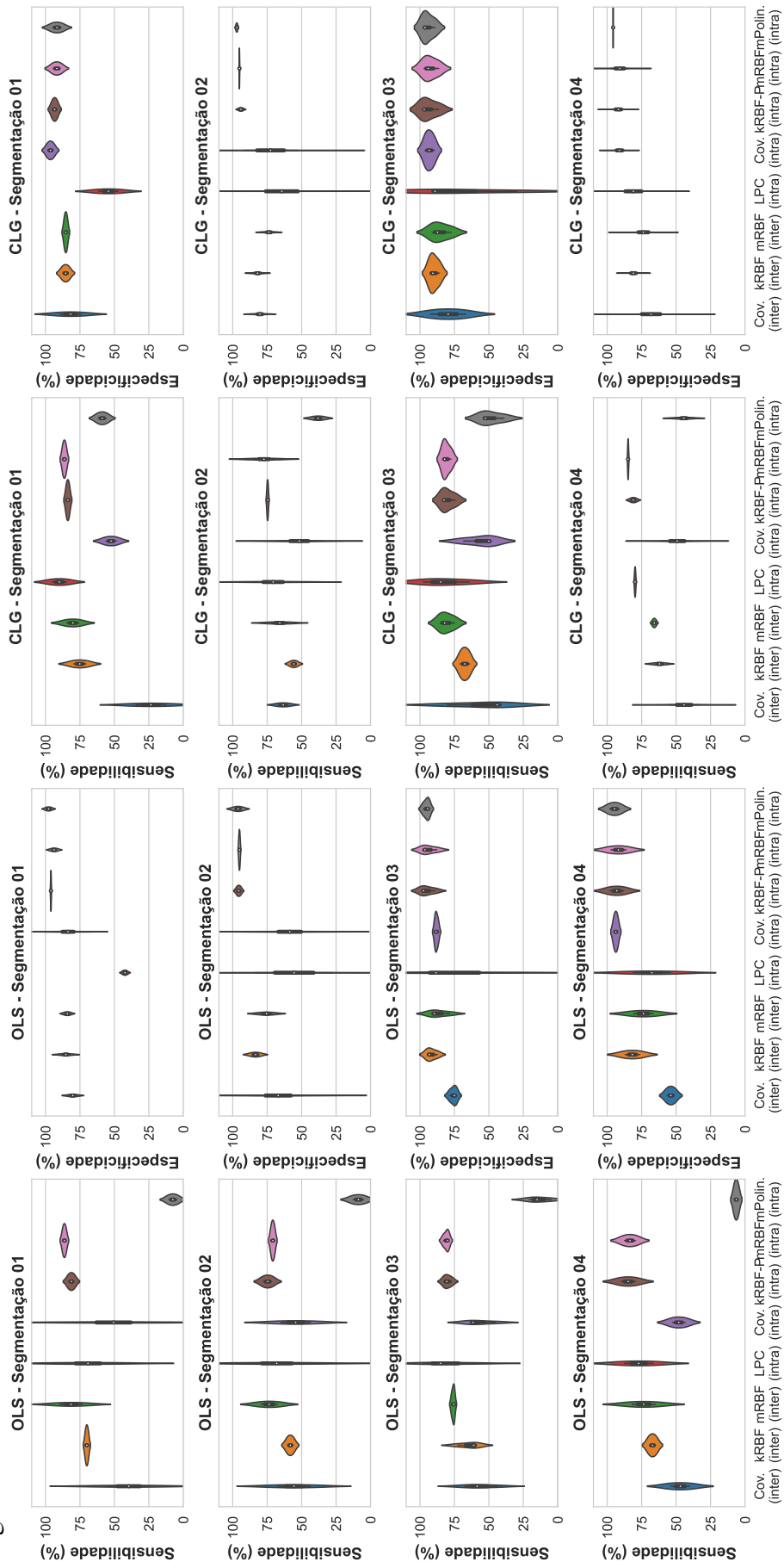
Fonte: Elaborado pelo Autor.

Figura 98 – Resultados do caso CHB17.



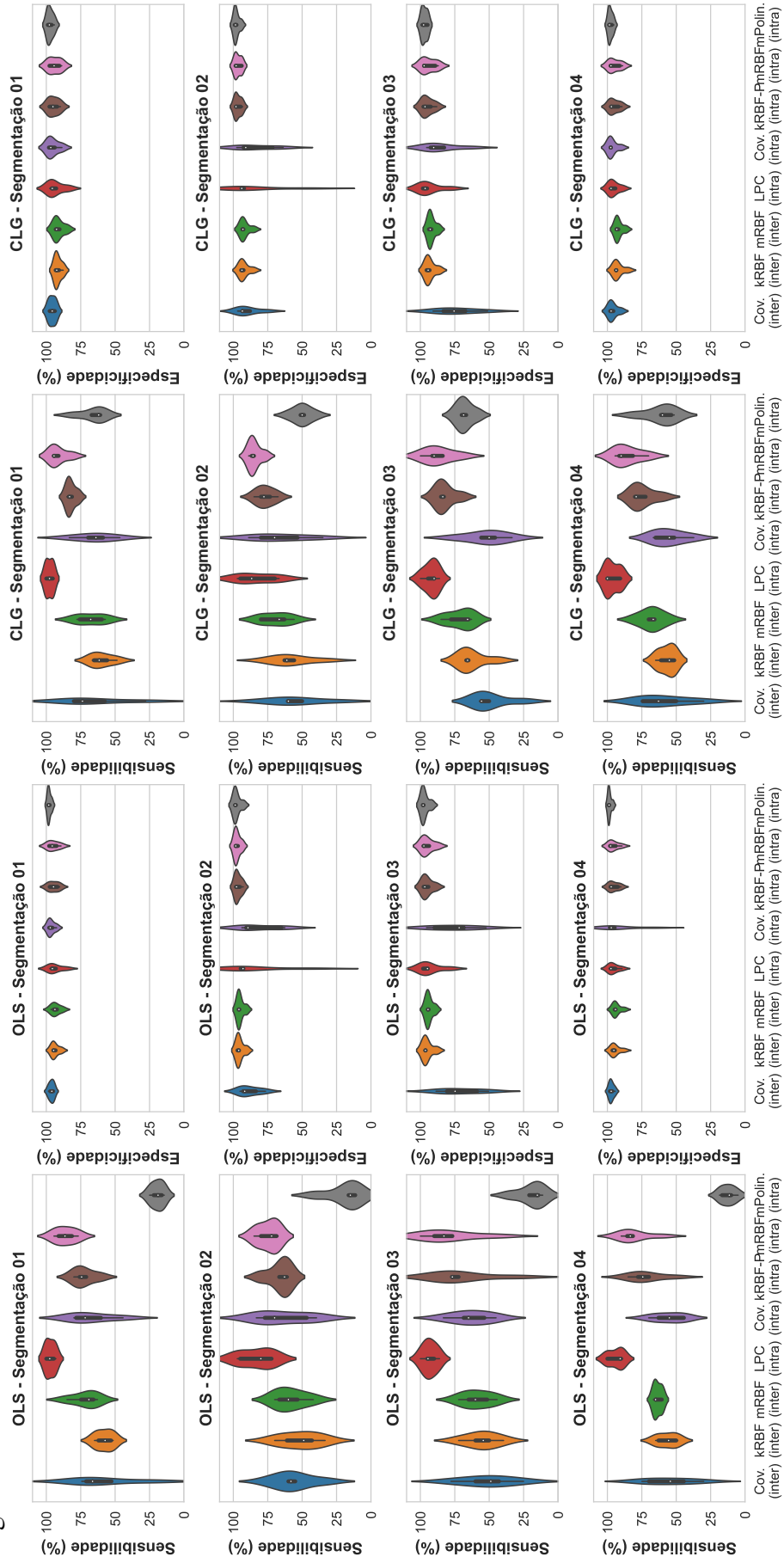
Fonte: Elaborado pelo Autor.

Figura 99 – Resultados do caso CHB19.



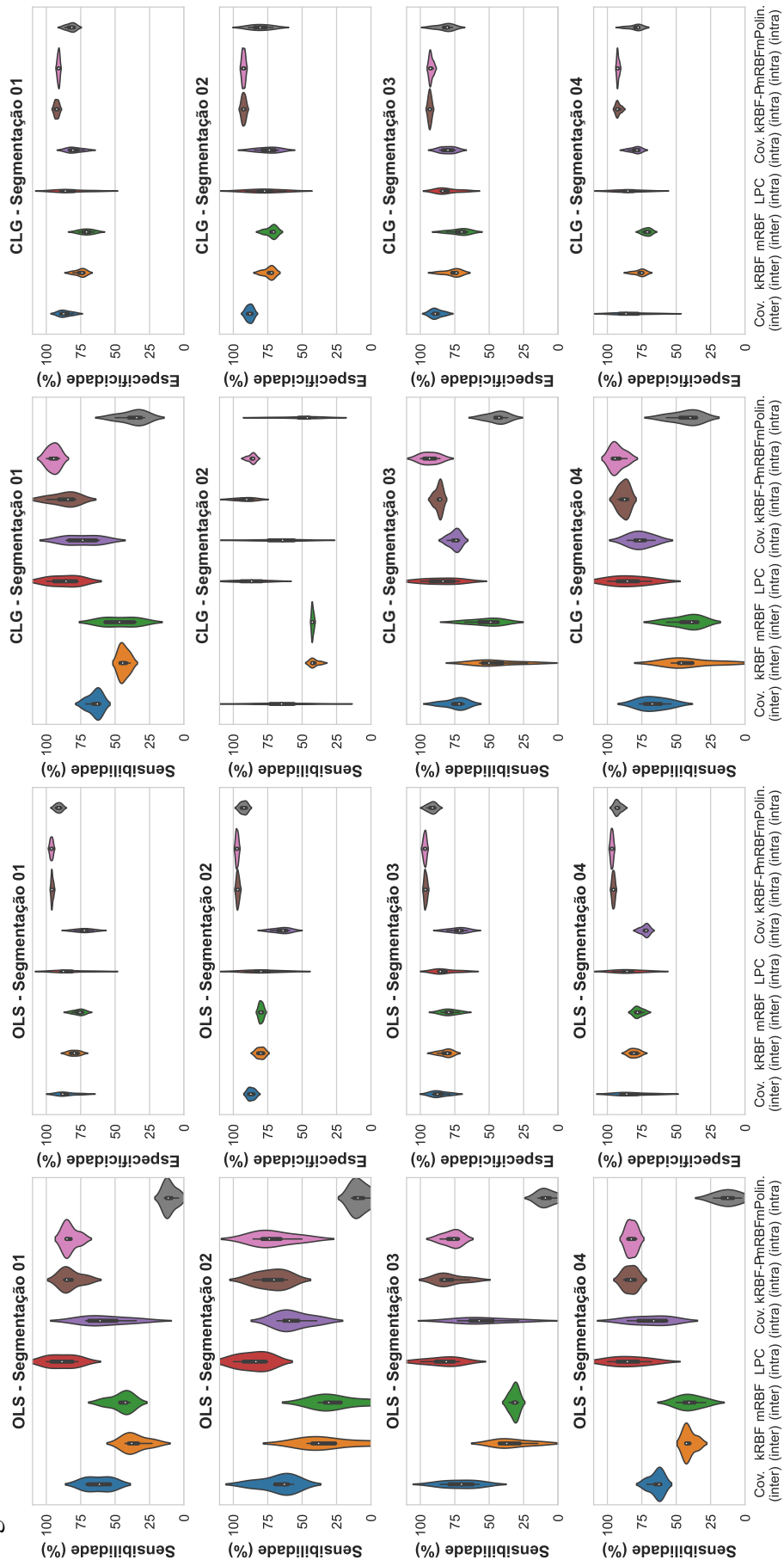
Fonte: Elaborado pelo Autor.

Figura 100 – Resultados do caso CHB20.



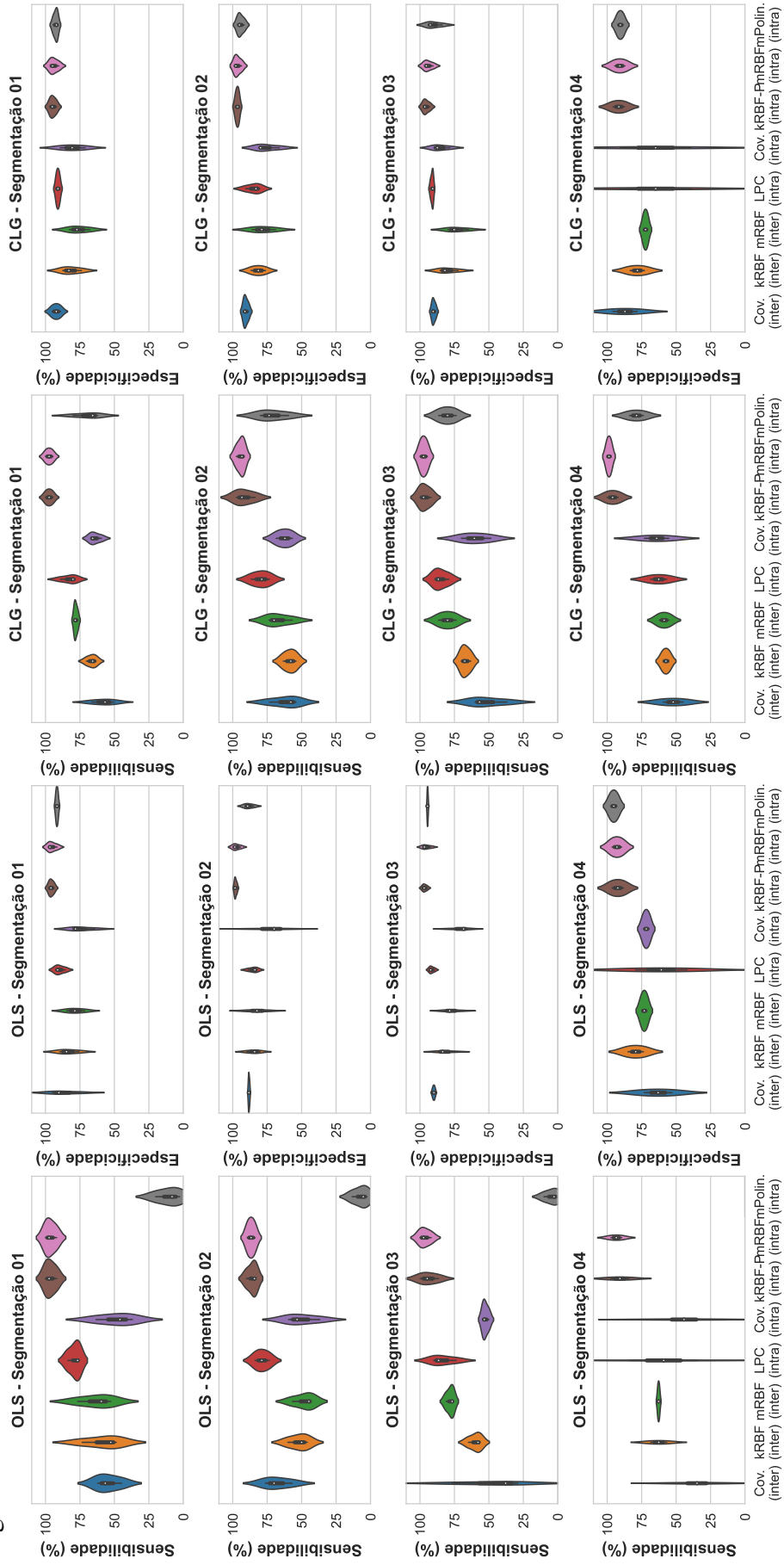
Fonte: Elaborado pelo Autor.

Figura 101 – Resultados do caso CHB21.



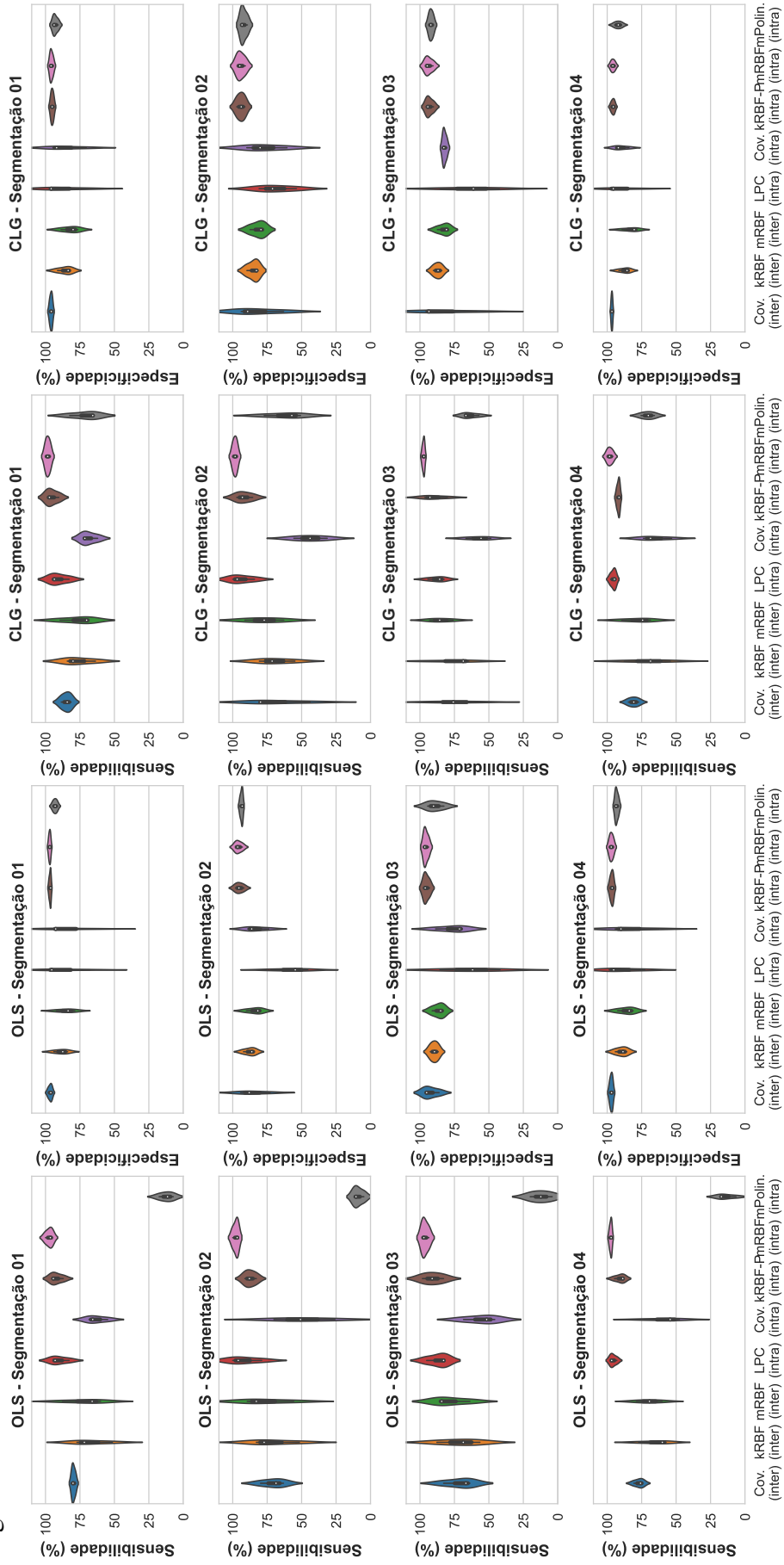
Fonte: Elaborado pelo Autor.

Figura 102 – Resultados do caso CHB22.



Fonte: Elaborado pelo Autor.

Figura 103 – Resultados do caso CHB23.



Fonte: Elaborado pelo Autor.