

Multilevel Mental Stress Detection Using Ultra-Short Pulse Rate Variability Series

Zubair e Yoon (2020)

1 Destaques

- Um sistema de detecção de estresse mental multinível é proposto usando séries PRV¹ (Pulse Rate Variability) ultracurtas;
- Um paradigma experimental que utiliza tarefas aritméticas mentais (MAT) é projetado para indução de estresse;
- Sinais PPG de ultracurto prazo foram registrados para estimativa de PRV;
- Um conjunto de recursos é proposto para capturar informações temporais de baixo nível no gráfico de Poincaré;
- A classificação de estresse em cinco níveis é realizada significativamente com SVM e QDC² (Quadratic Discriminant Classifier)

2 Resumo

Os autores propuseram um sistema de detecção de estresse multinível usando gravações de prazo ultracurto de um sensor vestível de baixo custo. Projetaram um paradigma experimental baseado em tarefas aritméticas mentais (MAT) para estimular adequadamente diferentes níveis de estresse. Durante o experimento, sinais de fotopletismograma (PPG) foram registrados junto com feedback subjetivo para validação da indução de estresse. As séries de intervalos batimento a batimento, estimados a partir de segmentos de sinais PPG de 60s de duração, foram usadas para extrair diferentes características com base em sua confiabilidade. Para capturar as informações temporais nos segmentos de prazo ultracurto de PPG, introduzimos um novo conjunto de características que tem o potencial de quantificar as informações temporais em nível ponto a ponto no gráfico de Poincaré. Também usamos um algoritmo Sequencial Forward Floating Selection (SFFS) para mitigar os problemas de irrelevância e redundância entre recursos. Investigamos dois classificadores baseados em análise discriminante quadrática (QDA) e Support Vector Machine (SVM). Os resultados do método proposto produziram 94,33% de

¹Variabilidade da Frequência de Pulso

²Classificador Discriminante Quadrático

precisão com SVM para identificação de cinco níveis de estresse mental. Além disso, validamos a generalização do sistema avaliando seu desempenho em um conjunto de dados registrado com um estressor diferente (Stroop). Em conclusão, descobrimos que o sistema de detecção de estresse multinível proposto em conjunto com novos parâmetros do gráfico de Poincaré tem o potencial de detectar cinco estados diferentes de estresse mental usando gravações de prazo ultracurto de um sensor PPG de baixo custo.

Palavras-chave: variabilidade da frequência de pulso, detecção de estresse mental, gráfico de Poincaré, sinais PPG, análise discriminante quadrática, máquina de vetor de suporte, sensores vestíveis.

3 Protocolos experimentais

3.1 Participantes

14 estudantes de pós-graduação (12 homens e 2 mulheres) foram recrutados para este experimento. Todos os participantes eram saudáveis e tinham entre 25 e 35 anos. Nenhum dos sujeitos sofria de qualquer doença crônica como patologia cardiovascular ou mental. A triagem dos participantes foi realizada usando um questionário de autorrelato após torná-los cientes do propósito e procedimento do experimento. No questionário de autorrelato, informações sobre duração do sono, consumo de álcool, café, tabaco e drogas foram coletadas e levadas em consideração antes da coleta de dados.

3.2 Método experimental

Dois experimentos diferentes para registrar sinais PPG relacionados ao estresse foram conduzidos de acordo com o código de ética da associação médica mundial. Ambos os experimentos usaram estressores visuais para provocar diferentes estados de estresse e foram projetados em Python usando PsychoPy (módulo Python de código aberto para neurociência e psicologia experimental). Os experimentos foram realizados em uma sala silenciosa para evitar perturbações e interferências externas. Ambos os experimentos começaram com uma fase de habituação de 5min para minimizar o efeito de qualquer estímulo recentemente experimentado. Depois, foi realizada uma fase de descanso de 5min, na qual os sujeitos foram solicitados a manter os olhos abertos enquanto estavam sentados em uma cadeira repousante. As figuras 1 e 2 ilustram os paradigmas experimentais supracitados.

3.2.1 Experimento 1

Um estímulo de estresse baseado em Tarefa Aritmética Mental (MAT) foi projetado usando o procedimento padrão. As tarefas foram divididas em quatro níveis com base no grau de dificuldade envolvendo adição, subtração, multiplicação e divisão de números até dois dígitos no máximo. Todas as tarefas foram projetadas para ter uma resposta de um único dígito. O tempo máximo de resposta de 5, 10, 15 e 20s foi definido para tarefas de nível 1, nível 2, nível 3 e nível 4, respectivamente.

- MAT Nível 1: $7+2$ ou $8-3$;
- MAT Nível 2: $5*2-8$ ou $3*2+3$;

- MAT Nível 3: $2-1*5+6$;
- MAT Nível 4: $10/5+2*2-5$.

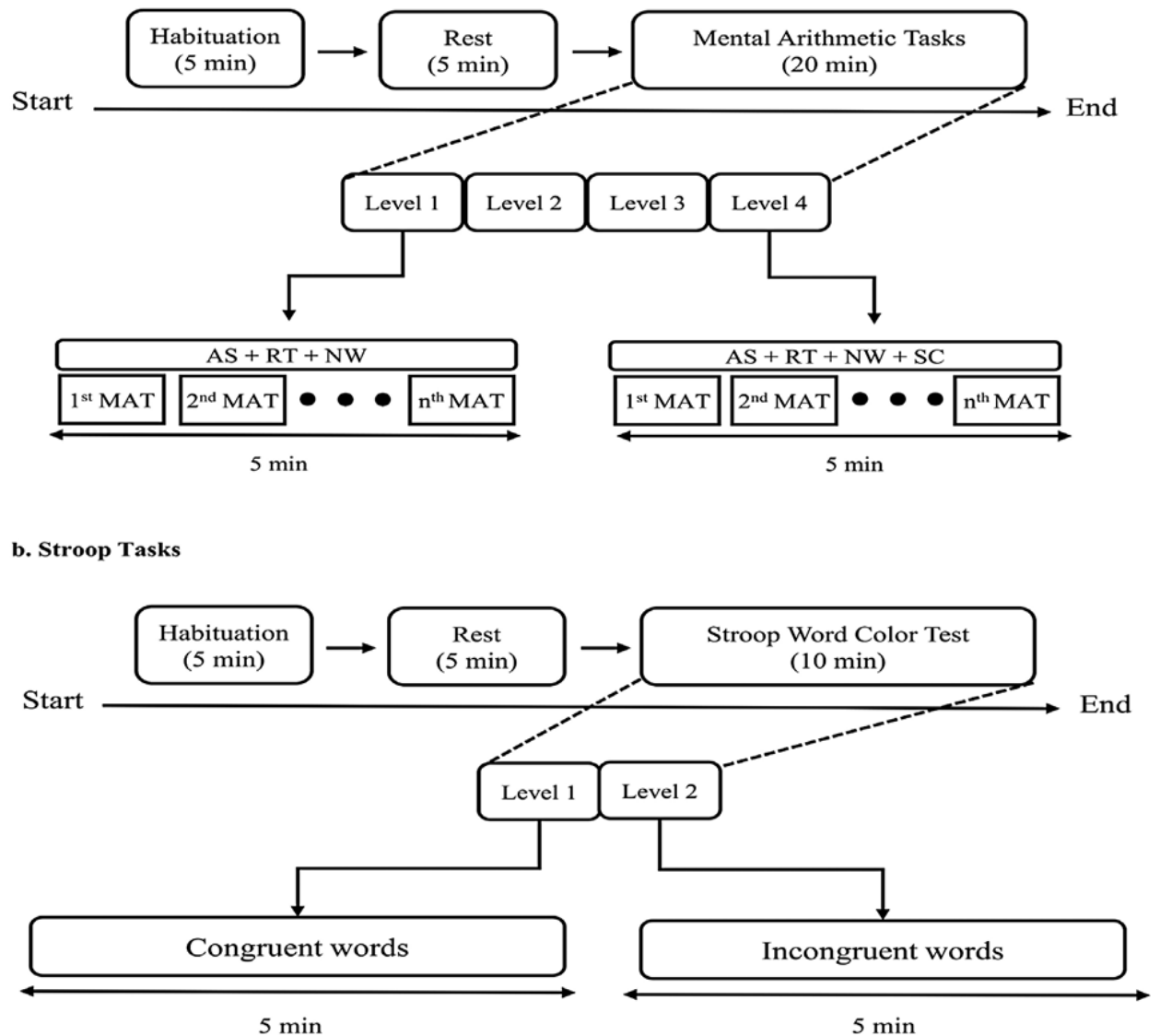


Figura 1: Tarefa Aritmética Mental (Zubair e Yoon, 2020). AS: Average Score, RT: Remaining Time, NW: Number of Wrong Answers, SC: Stressful Comments.

3.2.2 Experimento 2

O segundo experimento inclui dois níveis de Tarefas de Palavras Coloridas Stroop [37] em que uma única palavra de quatro nomes de cores diferentes (vermelho, amarelo, verde e azul) foi exibida aleatoriamente em uma tela de computador. O sujeito foi solicitado a digitar a primeira letra (r, y, g e b) da cor da fonte da palavra exibida. Durante o primeiro nível (Congruente) do teste Stroop, os nomes das cores foram exibidos em uma cor semelhante à palavra exibida (por exemplo, a palavra "azul" exibida na cor azul, enquanto no segundo nível (Incongruente),

os nomes das cores foram impressos em uma cor diferente da palavra impressa (por exemplo, a palavra "vermelho" exibida na cor verde).

O objetivo do teste de palavras congruentes e incongruentes é provocar diferentes níveis de estresse. Por exemplo, a palavra incongruente requer mais tempo e concentração do que o teste de palavras congruentes, induzindo assim um alto nível de estresse em comparação com a palavra congruente [31].

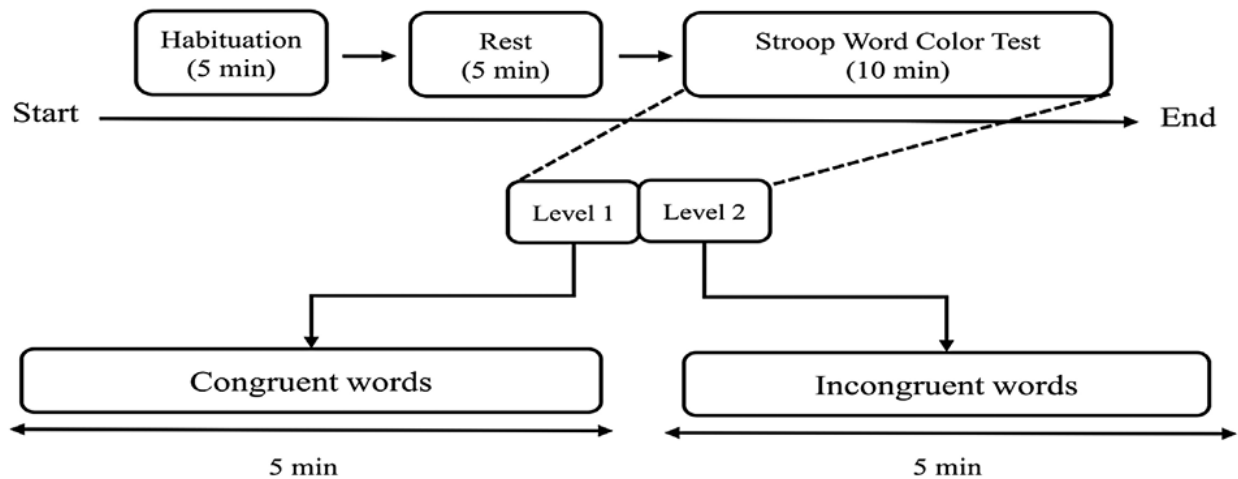


Figura 2: Teste Stroop (Zubair e Yoon, 2020).

Durante essas tarefas, após cada resposta, o sujeito recebia feedback (correto, incorreto) por meio de uma mensagem de exibição em um monitor de acordo com a resposta. A pontuação total, o número de respostas erradas e o tempo restante também eram exibidos na tela durante os experimentos. Nos níveis 3 e 4, para indução de estresse de nível mais alto, alguns comentários estressantes e negativos também foram exibidos em um certo número de trilhas. Esses comentários incluem: "seu desempenho está abaixo da média", "seu desempenho é decepcionante", "seu tempo de resposta é muito baixo", "não dê respostas erradas". A frequência desses comentários foi mantida mais alta no nível 4 para garantir a discriminação do estresse nos níveis 3 e 4.

Para validar a indução do estresse, juntamente com a pontuação de desempenho, também registramos o tempo de resposta de cada tarefa MAT e Stroop, trazidas através das Tabelas 2 e 3, respectivamente. Uma diminuição significativa na pontuação de desempenho e no tempo de resposta foi observada com o aumento do nível de dificuldade. Fatores como restrição de tempo, feedback de desempenho e comentários estressantes contribuíram muito para a indução de diferentes níveis de estresse.

Table 2
Participants score and response time at each level of MAT.

Stress level	Score (Mean $\pm$ SD)	Response time (ms) (Mean $\pm$ SD)
Level 1	62 $\pm$ 8	880 $\pm$ 120
Level 2	44 $\pm$ 12	6745 $\pm$ 3150
Level 3	40 $\pm$ 11	11,513 $\pm$ 1250
Level 4	38 $\pm$ 13	14,860 $\pm$ 1845

Table 3
Participants score and response time at each level of stroop.

Stress level	Score (Mean $\pm$ SD)	Response time (ms) (Mean $\pm$ SD)
Level 1	180 $\pm$ 6	680 $\pm$ 206
Level 2	95 $\pm$ 15	820 $\pm$ 180

3.3 Aquisição de dados

- Os sinais PPG foram adquiridos usando o Pulse Sensor Amped, que é projetado e fabricado pela World Famous Electronics LLC, Nova York, EUA;
- O Pulse Sensor Amped é incorporado com um LED verde e um sensor de luz. Ele computa as variações no volume sanguíneo capturando a luz refletida dos tecidos. As vantagens do vestível Pulse Sensor Amped são as que seguem listadas abaixo:
 - relação Custo-Benefício;
 - gravações de qualidade;
 - tamanho;
 - facilidade de uso
- O sensor PPG foi usado no dedo indicador da mão esquerda com uma tira de dedo para minimizar o ruído;
- Os sinais foram registrados da ponta do dedo usando uma placa Arduino;
- Os dados foram transferidos do Arduino para o computador usando comunicação serial;
- Durante o experimento, os dados do sensor foram registrados para cada nível separadamente, e seus respectivos arquivos de dados foram armazenados individualmente. Para esse propósito, incorporamos o programa de comunicação serial no programa MAT (PsychoPy) que lida com a duração de tempo de cada nível. Assim, durante cada nível, os dados seriais foram anexados às suas respectivas listas de dados. Este procedimento anula o uso de marcadores de limites separados e garante a sincronização dos dados;
- De um único sujeito, registramos 25min e 15min de dados PPG durante os experimentos MAT e Stroop, respectivamente. Considerando apenas as gravações PPG do experimento MAT, um total de 25 amostras por pessoa foram obtidas com uma janela não sobreposta de 60s. Da mesma forma, obtivemos um total de 350 amostras de 14 sujeitos, o que inclui 70 amostra de cada classe.

4 Metodologia

Um diagrama de blocos do sistema de reconhecimento de estresse proposto é ilustrado através da Figura 3. O sistema é composto por cinco etapas sequenciais básicas, a saber: a) Aquisição de dados, b) Pré-processamento, c) Extração de características, d) Seleção de características e e) Classificação.

4.1 Pré-processamento

Durante a gravação PPG os sinais são afetados por diferentes tipos de ruídos e artefatos que podem detrimir a precisão do sistema de detecção de estresse. O mais comum deles é o artefato de movimento, que tem grande influência na qualidade dos sinais PPG e pode levar a uma estimativa PRV ruim. Esse tipo de ruído geralmente é causado pelo movimento do sujeito ou pelo contato ruim entre a ponta do dedo e o sensor. Além do artefato de movimento, a interferência da linha de energia e a variação da linha de base induzida pela respiração e atividade muscular

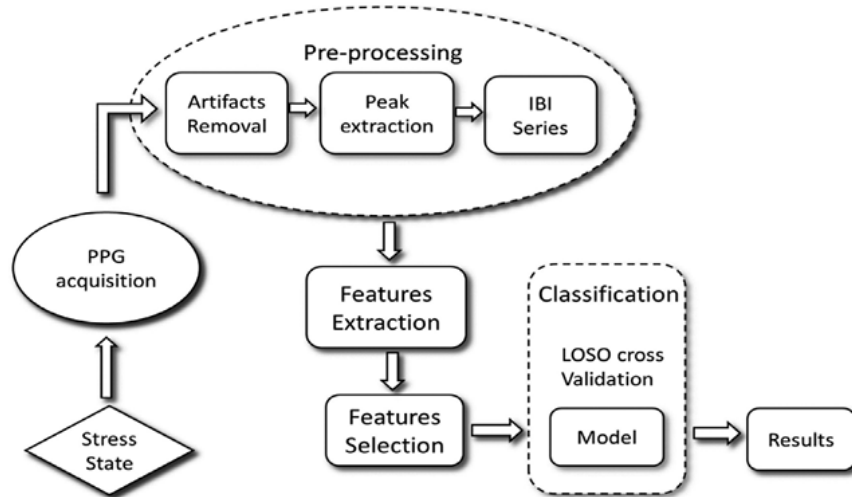


Figura 3: Diagrama de blocos do sistema de detecção de estresse (Zubair e Yoon, 2020).

também podem deteriorar a qualidade do sinal [47]. Para uma melhor estimativa da série PRV, Zbair e Yoon (2020), primeiro aplicaram um filtro passa-alta com uma frequência de corte de 0.5Hz para remover o desvio da linha de base. Depois, um filtro passa-baixa foi aplicado com uma frequência de corte de 35Hz para remover os ruídos de alta frequência e a interferência da linha de energia [47]. Após a redução de ruído do sinal PPG, estimamos os intervalos de tempo entre picos consecutivos. Para a detecção precisa de intervalos entre batimentos, a etapa essencial é identificar com precisão a localização do pico no sinal PPG. Para esse propósito, um algoritmo de posicionamento baseado em limiar foi adotado [50]. Uma ilustração visual do pré-processamento é dada através da Figura 4.

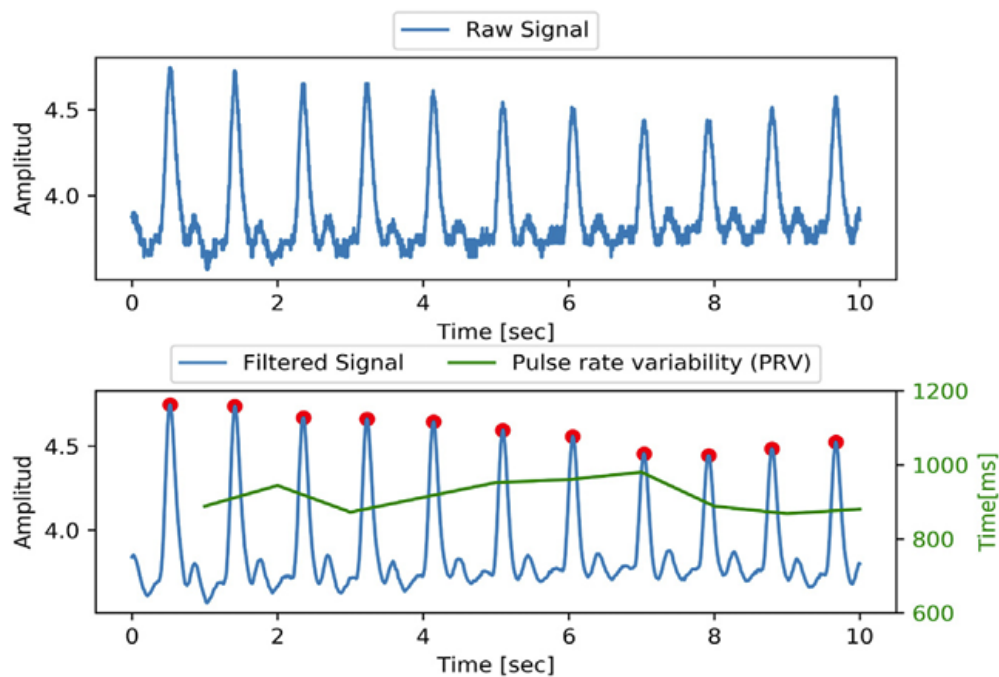


Figura 4: Pré-processamento do sinal PPG (Zubair e Yoon, 2020).

Embora as etapas iniciais de redução de ruído eliminem significativamente o ruído, incluindo artefatos de movimento e interferência de linha de energia, às vezes é difícil recuperar o sinal original devido ao alto nível de ruído e distorção. Estes sinais levam à estimativa errada de PRV e fornecem desempenho de classificação ruim. Para lidar com esses casos, consideramos apenas intervalos normais (com duração de tempo entre 600-1200ms). Mais especificamente, durante a estimativa de PRV, todos os intervalos anormais (intervalos com duração de tempo menor que 600ms ou maior que 1200ms) foram excluídos e apenas aqueles intervalos foram incluídos na faixa de 600 a 1200ms. No entanto, **excluir muitos intervalos anormais reduz o comprimento da série PRV resultante**. Portanto, segmentos PPG com intervalos anormais menores que 15% foi escolhido para garantir que, após a eliminação de intervalos anormais, o segmento PPG selecionado ainda tenha um comprimento de tempo maior que 50s. Para todos os participantes, os intervalos anormais em cada segmento PPG foram menores que 15% do total de intervalos e, portanto, nenhum segmento foi rejeitado com base nesse critério. Esta regra de eliminação foi seguida apenas para evitar a redução no comprimento de tempo dos dados, pois o comprimento dos dados afeta profundamente o desempenho geral do sistema. Todo esse processo garante a disponibilidade de séries PRV sem erros para detecção de nível de estresse.

4.2 Extração de recursos

Para analisar a variabilidade cardiovascular, medidas de HRV no domínio do tempo, domínio da frequência e não linear são comumente usadas. Neste estudo, diferentes medidas associadas ao estresse foram extraídas do PRV com base em sua confiabilidade para análise de HRV de prazo ultracurto.

4.2.1 Características padrão

- **Medidas de domínio de tempo:** Um conjunto de características de HRV de domínio de tempo padrão foi derivado de gravação de PPG de prazo ultracurto (60s). Calculamos a **média das séries temporais de PRV (RR média)**, **frequência cardíaca média (HR média)**, **desvio padrão das séries temporais de PRV (STDRR)** e **percentagem de intervalos batimento a batimento sucessivos que diferem em mais de 50ms (pNN50)**. **Da mesma forma, a raiz quadrada da diferença quadrática média de intervalos de batimento sucessivos (RMSSD) também foi calculada da forma que segue mostrado através da equação 1.**

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N-1} (RR_{j+1} - RR_j)^2} \quad (1)$$

onde:

RR: Intervalo de tempo entre duas batidas sucessivas e N é o número total de intervalos de batidas.

- **Medidas de domínio de frequência:** Medidas de domínio de frequência de HRV de curto prazo estimam a distribuição de potência em três bandas de frequência diferentes, a saber: a) banda de frequência muito baixa (VLF), b) banda de baixa frequência (LF) e c) banda de alta frequência (HF). As bandas VLF (0,003-0,04Hz) e LF (0,04-0,15Hz)

são moduladas pelo sistema nervoso simpático e parassimpático. No trabalho de Zubair e Yoon (2020), as características do domínio de frequência computadas incluem potência HF, potência LF, potência LF, LF/HF (razão de potência LF para HF), norma LF (potência relativa da banda LF) e norma HF (potência relativa da banda HF). Zubair e Yoon (2020) empregaram o método do Periodograma de Welch baseada na Transformada Rápida de Fourier (FFT) para análise espectral de potência de diferentes bandas de frequência. Para todos os segmentos do PRV, a janela "Hanning" foi usada com uma sobreposição máxima de 90%.

- **Medidas não lineares:** Medidas não lineares de HRV incluem entropia de amostra, entropia aproximada, análise de flutuação sem tendência, análise de gráfico de Poicaré e análise de gráfico de recorrência. No entanto, neste trabalho, incluímos apenas parâmetros do gráfico de Poincaré, pois a confiabilidade desses parâmetros para HRV de prazo ultra-curto foi relatada em [48]. O gráfico de Poincaré é um método geométrico e não linear para analisar a HRV e representa a correlação entre intervalos inter-batimentos sucessivos. Ele é gerado plotando o intervalo batimento a batimento em relação a um intervalo anterior para formar um gráfico disperso, conforme apresentado na Figura 5 abaixo.

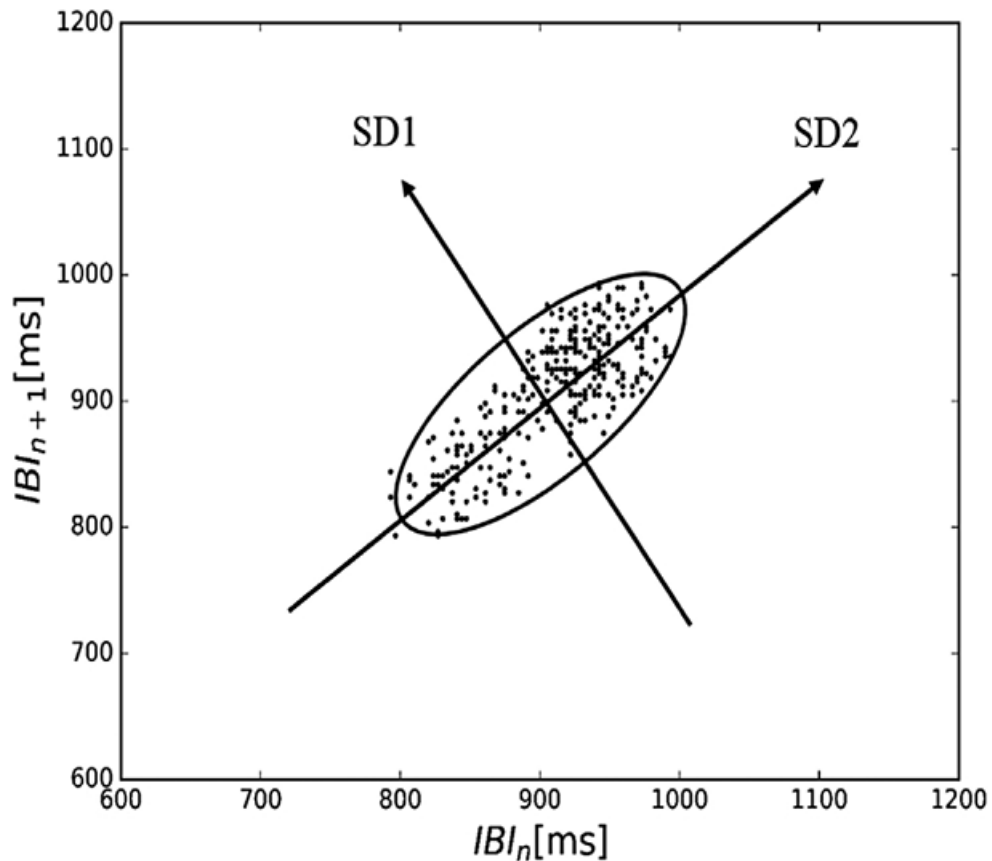


Figura 5: Gráfico de Poincaré de intervalos entre batimentos (5min) com descritores padrão (SD1, SD2) (Zubair e Yoon, 2020).

Para análise quantitativo do gráfico de Poincaré, o descritor padrão (SD1 e SD2) pode ser obtido ajustando uma elipse ao gráfico. SD1 é o desvio padrão da variabilidade do intervalo inter-batimento de curto prazo que ilustra a dinâmica de curto prazo da HRV e pode ser computado como trazido através da equação 2 abaixo.

$$SD1 = \frac{\sqrt{2}}{2} std(x_i - x_{i+1}) \quad (2)$$

onde:

sd é o desvio padrão da série temporal da HRV e x representa o intervalo de tempo entre batimentos sucessivos. Da mesma forma, SD2 é o desvio padrão da variabilidade do intervalo entre batimentos de longo prazo que descreve a dinâmica de longo prazo da HRV e pode ser computado como trazido através da equação 3 abaixo.

$$SD2 = \sqrt{2std(x_i)^2 - \frac{1}{2}std(x_i - x_{i+1})^2} \quad (3)$$

Além de SD1 e SD2, também calculamos SD12 ($SD12 = SD1/SD2$) e S ($S = \pi SD1 \times SD2$). SD12 é a razão entre SD1 e SD2 que determina a aleatoriedade em HRV, enquanto S é a área sob a elipse e determina a variabilidade total do gráfico de Poincaré.

4.2.2 Características propostas

Zubair e Yoon (2020) introduzem um novo recurso: Medida de Transição de Ponto (PTM - Point Transition Measure) para quantificar a variação temporal no nível ponto a ponto do gráfico de Poincaré. Embora os descritores padrão (SD1 e SD2) do gráfico de Poincaré tenham um bom desempenho, no entanto, esses descritores carregam apenas informações espaciais e não incorporam a variação temporal. Por exemplo, é possível que os sinais de HRV tenham distribuição de pontos idêntica com valores SD1 e SD2 semelhantes, mas estruturas temporais diferentes. Tal caso é descrito na Figura 6 que é trazida abaixo.

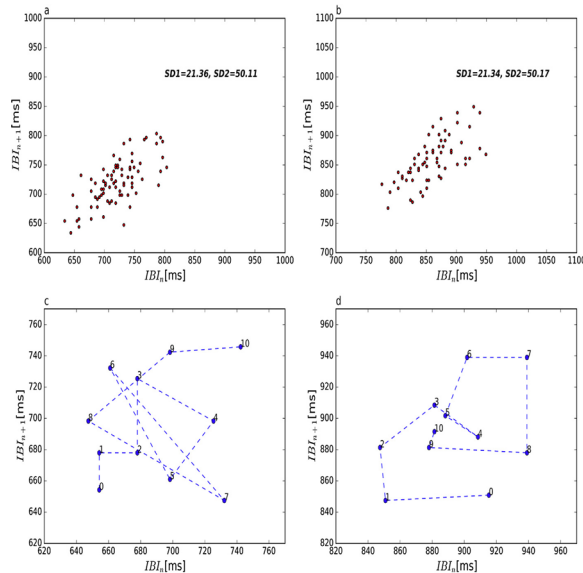


Figura 6: Painel superior (a,b): Gráfico de Poincaré de duas séries temporais com $N = 70$ e quase os mesmos DP1 e DP2. Painel inferior (c,d): Dinâmica temporal dos primeiros 11 pontos de ambas as séries temporais. (Zubair e Yoon, 2020).

4.3 Análise estatística

A análise estatística foi realizada para examinar o potencial de características derivadas de segmentos PRV para discriminar entre cinco níveis de estresse (relax, L1, L2, L3 e L4). Para esse propósito, empregamos o teste ANOVA (Análise de Variância), unidirecional usando Python para determinar mudanças significativas entre características PRV derivadas de cinco grupos diferentes. Um *valor-p* menor que 0,05 foi considerado estatisticamente significativo [25, 30, 33].

4.4 Seleção de recursos

Após computar as características para todos os segmentos, a padronização dos dados foi realizada usando a técnica de padronização z-score para transformar os dados em uma escala comum com média zero e desvio padrão um, através da computação trazida através da equação 8 abaixo.

$$x_{new} = \frac{x - \mu}{\sigma} \quad (8)$$

onde x denota uma única característica informativa, μ e σ são a média e o desvio padrão, respectivamente. Depois, realizamos a seleção de características para eliminar as características irrelevantes (características com baixa correlação característica-classe) e redundantes (características com alta correlação característica-característica) a fim de mitigar o problema de deterioração do desempenho devido à maldição da dimensionalidade. O algoritmo *Sequential Floating Forward Selection* (SFFS) foi empregado para atingir o objetivo da seleção de características [49]. No estudo de Zubair e Yoon (2020), um classificador discriminante quadrático (QDC) foi usado como um estimador para SFFS como técnica de validação cruzada de 10 dobras. Conjunto de características superior (1-20) foram usados para treinar o classificador e o conjunto com maior desempenho de classificação e o número mínimo de características foi selecionado.

4.5 Classificação

Para classificação de características selecionadas, investigamos o desempenho de dois algoritmos de classificação: a) Análise Discriminante Quadrática (QDA) e b) Máquina de Vetores de Suporte (SVM). O classificador SVM foi empregado porque tem o potencial de lidar com dados não lineares. Além disso, os modelos SVM têm maior estabilidade e generalização e, portanto, são menos afetados por sobre-ajustes. O QDA foi empregado para fins de comparação, pois os modelos baseados em QDC também são menos propensos a sobreajuste.

Em ambos os métodos, as entradas do classificador foram as características selecionadas baseadas em SFFS que foram extraídas de segmentos PRV de 60s de comprimento, enquanto a saída do classificador foi um único nível de estresse. Para determinar a eficácia do modelo quantitativamente, usamos três parâmetros que incluem *sensibilidade*, *especificidade* e *precisão*. Antes de calcular esses parâmetros, obtivemos as variáveis abaixo:

- Verdadeiro Positivo (TP - True Positive): o número de casos que são identificados com precisão como casos;

- Falso Positivo (FP - False Positive): o número de casos que são identificados incorretamente como não casos;
- Verdadeiro Negativo (TN - True Negative): o número de não casos que são identificados com precisão como não casos;
- Falso Negativo (FN - False Negative): o número de não casos que são identificados incorretamente como casos.

Depois, calculamos a sensibilidade (taxa de verdadeiro positivo) e a especificidade (taxa de verdadeiro negativo) usando as equações 9 e 10 abaixo, respectivamente.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (9)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (10)$$

Além da sensibilidade e especificidade, também calculamos a precisão usando a equação 11, que demonstra a capacidade geral do modelo de diferenciar diferentes estados de estresse.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

Para avaliar o desempenho de classificação dos modelos, usamos o procedimento Leave-One-Subject-Out (LOSO). Nessa técnica, um sujeito (conjunto de teste) foi excluído e os dados dos treze sujeitos restantes (conjunto de treinamento) foram usados para treinar um classificador. O modelo treinado foi então avaliado no grupo de retenção (conjunto de teste) e as medidas de desempenho (precisão, especificidade e sensibilidade) foram computadas. Esse processo foi repetido 14 vezes, e cada vez um sujeito único foi usado como um conjunto de teste. Os resultados finais foram obtidos pela média das medidas de desempenho de 14 classificadores. Adotamos o método de validação cruzada LOSO porque ele reduz significativamente o viés e oferece melhor capacidade de generalização do que a divisão simples de treinamento/teste.

5 Resultados e discussões

5.1 Efeitos do estresse nas medidas de HRV

Quatro grupos diferentes de tarefas aritméticas com diferentes níveis de dificuldades foram usados para induzir quatro níveis de estresse. A indução de estresse foi confirmada pela análise do tempo de resposta registrado e da pontuação de desempenho dos sujeitos em cada nível do experimento. Por exemplo, o desempenho da tarefa e o tempo de resposta dos sujeitos foram significativamente reduzidos de baixo nível de estresse para alto nível de estresse, conforme apresentado nas Tabelas 2 e 3. Além disso, variações nos valores das características também foram examinadas para validar a indução de níveis de estresse.

Os sinais PPG coletados durante cada nível do experimento foram divididos em segmentos curtos de 60s. Após estimar o PRV, a extração de características da série PRV de prazo ultra-curto foi realizada computando o domínio do tempo, o domínio da frequência e medidas não

lineares e quase todas as características mostraram mudanças notáveis durante diferentes níveis de estresse. Na Tabela 4, resumimos a média e o desvio padrão para todas as características extraídas de séries temporais PRV de 60s de duração durante cada sessão do experimento para demonstrar como essas características são influenciadas pelo estresse mental. Muitos especialistas agumentaram que a média de RR, STDRR, RMSSD e pNN50 diminuem notavelmente de baixo estresse para alto estresse, enquanto o valor da média HR aumentou do estado de baixo estresse (L1) para o estado de alto estresse (L4). No entanto, observamos que o valor de STDRR aumentou de nenhum estresse para alto estresse. Embora esse comportamento da característica STDRR esteja em contradição com [31], ele corresponde bem com [46].

A análise estatística do domínio do tempo, domínio da frequência e medidas não colineares indica que esses parâmetros são significativamente diferentes entre diferentes estados de estresse. As variações nas medidas do domínio do tempo de HRV durante um curto período (60s) refletem as rápidas mudanças no sistema nervoso autônomo. Da mesma forma, diferenças significativas nas medidas no domínio da frequência entre diferentes estados também foram observadas. A norma LF, LF/HF e LF aumentaram de relaxado para alto estado de estresse, enquanto HF e norma HF diminuíram de nenhum estado de estresse para alto estado de estresse (L4). Quanto às medidas não lineares de PRV, diferenças significativas também foram observadas em SD1, SD2, SD12 e S entre diferentes níveis de estresse. O novo conjunto proposto de características também demonstrou mudanças significativas em todos os cinco níveis de estresse. Os valores de PTM, PTM+ e PTMnorm aumentaram de nenhum estresse para alto estresse. Em suma, as variações nas características extraídas de HRV foram consideradas consistentes com alterações no sistema nervoso autônomo e componentes para discriminar diferentes níveis de estresse.

5.2 Classificação do estresse multinível

Após a extração de características, a seleção de características foi realizada e um conjunto de características ótimo foi selecionado usando o algoritmo SFFS. Nove características mais relevantes e não redundantes foram selecionadas, que incluem RR médio, HR médio, RMSSD, Potência total, Lfnu, SD1, SD2, SD12 e PTM+. Para a seleção de características, o QDC foi empregado como um estimador no algoritmo SFFS. As precisões de diferentes números de características selecionadas são apresentadas na Figura 7 abaixo.

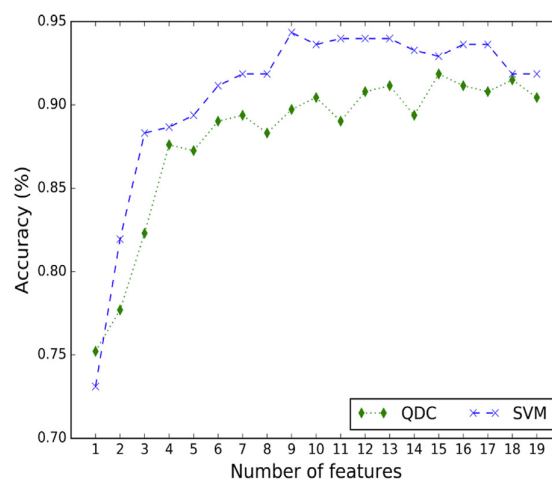


Figura 7: (Desempenho de QDC e SVM com diferentes números de recursos selecionados (Zubair e Yoon, 2020).

O conjunto selecionado de nove características foi usado para classificar cinco classes diferentes (repouso, L1, L2, L3, L4) usando os classificadores QDC e SVM. A Tabela 5 mostra o desempenho geral da classificação de ambos os classificadores com base em parâmetros selecionados de PRV ultracurto adquiridos durante o teste MAT. Em comparação com o QDC, descobrimos que o classificador SVM teve melhor desempenho no reconhecimento de cinco níveis diferentes de estresse com o conjunto selecionado de nove características.

Classifier	Levels	Sensitivity	Specificity	Accuracy
SVM	Relax	89.65%	97.61%	95.57%
	L1	77.77%	97.89%	94.69%
	L2	76.92%	94.25%	90.26%
	L3	100%	92.55%	93.38%
	L4	85.71%	100%	97.34%
	Total	86.01%	96.46%	94.33%
QDC	Relax	51.72%	98.80%	86.72%
	L1	77.77%	94.73%	92.03%
	L2	88.46%	90.80%	90.26%
	L3	73.68%	92.55%	89.38%
	L4	85.71%	91.30%	90.26%
	Total	75.46%	93.63%	89.73%

Figura 8: (Resultados da classificação de QDC e SVM para MAT (Zubair e Yoon, 2020).

As precisões gerais de classificação para SVM e QDC foram 94,33% (sensibilidade 86,01%, especificidade 96,46%) e 89,73% (sensibilidade 75,46%, especificidade 93,63%), respectivamente. Entre todas as cinco classes, amostras de dados do estado de estresse mais alto (L4) foram classificadas com pontuação de precisão de 97,34%, o que aprova a significância dos elementos evocadores de estresse no paradigma experimental. O uso de elementos distrativos e evocadores de estresse, como comentários estressantes, pressão de tempo e monitoramento de desempenho, contribuíram bem para induzir o estado de estresse mais alto e revelar a reação verdadeira e mais distinta do sistema nervoso autônomo.

5.3 Especificidade do estressor

Para validar que o desempenho do reconhecimento do modelo treinado não está relacionado a um único estressor, testamos o modelo treinado em dados gerados com um estressor diferente. Para esse fim, foi realizado um teste Stroop de três níveis e os dados dos três níveis foram anotados com os rótulos semelhantes aos três primeiros níveis do teste MAT. A Figura 9 mostra a comparação do desempenho da classificação entre QDC e SVM para três níveis de teste de Stroop. Durante este teste, uma pontuação média de classificação de 81,66% e 86% foi registrada para QDC e SVM, respectivamente. Os resultados gerais do teste de Stroop foram satisfatório e confirmam a capacidade de generalização do nosso modelo, no entanto, a pontuação de classificação por classe foi baixa, no entanto, a pontuação de classificação por classe foi baixa em comparação com as tarefas Stroop [51].

5.4 Comparação de desempenho

Finalmente, comparamos o desempenho do nosso sistema de detecção de estresse com outros estudos. No entanto, deve-se notar, nesse domínio, uma comparação justa entre diferentes estudos é difícil devido à falta de padronização nos conjuntos de dados. Os dados fisiológicos necessários para o treinamento de um modelo de detecção de estresse depende muito do tipo de estressor, da estratégia de utilização do estressor e do dispositivo usado para gravar dados. Os autores tentaram comparar o desempenho dos modelos de detecção de estresse com base no número de classes e no comprimento do segmento. A comparação do trabalho proposto com estudos anteriores ilustra que o sistema de detecção de tensão proposto alcançou maior escore de identificação do que outros métodos.

6 Conclusões e trabalhos futuros

Este artigo forneceu um sistema de detecção de tensão multinível usando gravações de termo ultracurto de um sensor PPG de baixo custo. Um modelo experimental baseado no teste MAT foi usado com diferentes elementos distritos e evocadores de estresse (restrição de tempo, feedback de desempenho e comentários estressantes) para uma indução eficiente da tensão multinível. A estimativa da série de variabilidade de intervalo batida a batimento foi realizada usando sessenta segundos de segmentos de sinais de PPG. Um conjunto de novos recursos de plotagem de Poincaré foi usado para capturar as informações temporais relacionadas ao estresse incorporadas em segmentos curtos de PPG. Ao analisar os resultados experimentais, descobrimos que o sistema de detecção de tensão proposto é capaz de detectar cinco níveis diferentes de estresse efetivamente com uma precisão de 94,33%. Devido à capacidade de identificação de tensão multinível, o sistema profissional também pode facilitar os médicos a monitorar o estresse mental de maneira eficaz. O trabalho futuro inclui a integração de diferentes sinais fisiológicos para avaliar o comportamento humano, juntamente com a detecção de estresse de vários níveis.