

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Sono ao Volante – Machine Learning para Previsão e Detecção de Sonolência

Cláudia Catarina Carvalho Rodrigues



Mestrado Integrado em Engenharia Informática e Computação

Orientador: Professor Luís Paulo Reis

Co-orientadora: Professora Brígida Mónica Faria

31 de Julho de 2020

Sono ao Volante – Machine Learning para Previsão e Detecção de Sonolência

Cláudia Catarina Carvalho Rodrigues

Mestrado Integrado em Engenharia Informática e Computação

Resumo

O sono insuficiente é um problema proeminente na sociedade moderna, com diversos efeitos e riscos negativos. Uma das consequências mais graves são os acidentes de viação causados pela sonolência, que prejudica a percepção, o tempo de reação e todo o processo de decisão do condutor. Os condutores profissionais são um dos grupos de risco para estes acidentes, nomeadamente por privação de sono, condução noturna, longas distâncias a percorrer e distúrbios de sono.

As soluções atuais para combater o sono ao volante abordam o problema pela deteção de sonolência. Assim, o condutor precisa de chegar a esse estado para ser emitido um alerta, o que pode ser demasiado tarde e este já não ter reflexos para reagir. Neste contexto, torna-se relevante o desenvolvimento de um sistema que integre a previsão de sonolência e a sua prevenção. Através da previsão do estado de sono, será possível alertar antecipadamente o condutor enquanto este ainda se encontra em vigília. Para minimizar a ocorrência de novos episódios, é necessário identificar o motivo pelo qual ocorre um estado de sonolência indevidamente, causado por distúrbio ou privação do sono. Com vista numa solução não intrusiva e de baixo custo, a sua operacionalização será através da análise de dados biométricos de um dispositivo de pulso utilizado em permanência.

Este trabalho está integrado no projeto “Sono Ao Volante 2.0” e foram várias as contribuições, como seja: a revisão da literatura associada aos conceitos do sono e distúrbios do sono; apresentação das principais soluções científicas e comerciais para recolha de dados; análise do sinal de electrocardiograma para a classificação automática das fases do sono. Para tal, foi utilizado um *dataset* público, e extraídos atributos associados à variabilidade da frequência cardíaca. Na classificação de três e quatro classes, o algoritmo *Random Forest* permitiu atingir uma exatidão de 90% e 78% e um coeficiente Kappa de Cohen de 0.71 e 0.64, respetivamente. Apesar de promissora, algumas limitações ainda estão presentes nesta abordagem, nomeadamente, a viabilidade de um cenário dependente do indivíduo, sendo necessário um conjunto de dados mais abrangente para análise futura.

Palavras-chave: Previsão de sonolência, Sinais biométricos, Sistema não intrusivo, *Machine learning*.

Abstract

Insufficient sleep is a prominent problem in the modern society with several negative effects and risks. One of the most serious consequences is traffic accidents caused by drowsiness, which impairs the driver's perception, reaction time and decision-making. Professional drivers are particularly at risk of such accidents, namely due to sleep deprivation, night-time driving, long distances to travel and sleep disorders.

The current solutions to prevent a person from falling asleep while driving address the problem by detecting drowsiness. Thus, the driver must reach this state to receive an alert, which may be too late, and he can no longer react. In this context, it is relevant to develop a system that integrates the prediction of drowsiness and its prevention. By predicting the drowsy state, it will be possible to warn the driver in advance while he is still alert. To minimize the occurrence of further incidents, there is a need to identify the reason why a state of drowsiness occurs, caused by a disorder or sleep deprivation. Towards a non-intrusive and low-cost solution, the system will be implemented through the physiological signals collected from a wearable device used continuously.

This work is integrated on the project "Sono Ao Volante 2.0" and there were several contributions, such as: a literature review of the concepts of sleep and sleep disorders; presentation of the main scientific and commercial solutions for data collection; analysis of electrocardiogram signal for automatic sleep staging. For that purpose, a public dataset was used, and features related to heart rate variability were extracted. In the three and four class classification problem, the algorithm Random Forest reached an accuracy of 90% and 78% and a Cohen's kappa statistic of 0.71 and 0.64, respectively. Although promising, some limitations are still present in this approach, namely, the viability of a user dependent scenario, requiring a larger dataset in future analysis.

Keywords: Drowsiness prediction, Biometric data, Non-intrusive system, Machine learning.

Agradecimentos

Gostaria de agradecer aos meus orientadores, Professor Luís Reis e Professora Brígida Faria, pelo apoio e oportunidade de participar neste projeto; à Faculdade pelo conhecimento transmitido que me motivou a seguir este caminho; aos meus colegas por todos os momentos proporcionados ao longo da minha vida académica; e à minha família, por estar sempre comigo.

Cláudia Rodrigues

*“Tudo o que a mente humana pode conceber,
ela pode conquistar.”*

Napoleon Hill

Conteúdo

1	Introdução	1
1.1	Contexto	1
1.2	Motivação	2
1.3	Objetivos	2
1.4	Contribuição	3
1.5	Estrutura do documento	3
2	Revisão da literatura	5
2.1	Sono e ritmo circadiano	5
2.2	Padrões de sono	7
2.3	Métodos para detecção de sonolência	9
2.4	Análise de sinais fisiológicos	11
2.4.1	Eletroencefalograma	12
2.4.2	Eletrocardiograma	12
2.4.3	Eletro-oculograma	13
2.4.4	Eletromiograma	14
2.4.5	Fotopletismografia	14
2.4.6	Propriedades da pele	14
2.5	Deteção de sonolência	15
2.5.1	Soluções comerciais	17
2.6	Previsão da sonolência	17
2.7	Classificação do sono	18
2.7.1	Soluções comerciais	21
2.8	Deteção de apneia do sono	21
2.8.1	Soluções comerciais	24
2.9	Sensores e Datasets	24
2.9.1	Datasets	24
2.9.2	Sensores	25
2.9.3	Discussão	26
3	Classificação do sono	27
3.1	Dados	27
3.1.1	Pré-processamento	27
3.1.2	Extração de atributos	28
3.2	Algoritmos	30
3.2.1	Support Vector Machine	30
3.2.2	Análise Discriminante Linear	30
3.2.3	Random Forest	31

3.2.4	K-Nearest Neighbors	32
3.2.5	Rede Neuronal Recorrente	32
3.3	Validação	33
3.3.1	Validação cruzada	33
3.3.2	Dados não balanceados	33
3.3.3	Seleção de atributos	35
3.3.4	Medidas de desempenho	35
3.4	Resultados	37
3.4.1	Resultados da classificação	37
3.4.2	Atributos selecionados	41
3.4.3	Análise sequencial	42
3.4.4	Discussão e limitações	43
4	Conclusões e trabalho futuro	45
4.1	Conclusões	45
4.2	Trabalho futuro	46
	Referências	47

Lista de Figuras

1.1	Arquitetura geral da solução.	3
2.1	Modelo de regulação do sono.	6
2.2	Arquitetura do sono.	7
2.3	Leituras poligráficas das fases do sono.	8
2.4	Escala de sonolência de Karolinska.	9
2.5	Deteção de sonolência no condutor baseada em sinais.	11
2.6	Partição do espectro de potência da variabilidade da frequência cardíaca.	13
2.7	Análise de batimentos cardíacos com o eletrocardiograma e fotopletismografia.	14
2.8	Morfologia do eletrocardiograma.	22
3.1	Abordagem para processamento dos segmentos.	28
3.2	Exemplo de um segmento do eletrocardiograma.	28
3.3	Séries obtidas do eletrocardiograma	29
3.4	Hiperplano de classificação do algoritmo <i>Support Vector Machine</i>	30
3.5	Aplicação da análise discriminante linear.	31
3.6	Arquitetura do algoritmo <i>Random Forest</i>	31
3.7	Aplicação do algoritmo <i>K-Nearest Neighbors</i>	32
3.8	Arquitetura básica de uma rede neuronal recorrente.	32
3.9	Validação cruzada <i>k-fold</i>	33
3.10	Aplicação de sub-amostragem e sobre-amostragem.	34
3.11	Aplicação do algoritmo SMOTE.	34
3.12	Comparação de curvas ROC com diferentes valores de área sob a curva.	36
3.13	Ilustração da abordagem utilizada.	37
3.14	Arquitetura de redes com memória bidirecional de longo prazo.	42

Lista de Tabelas

2.1	Comparação entre medidas para detecção de sonolência.	10
2.2	Estudos de detecção de sonolência no condutor.	16
2.3	Estudos de classificação do sono.	20
2.4	Exemplos de monitores de sono.	21
2.5	Estudos de detecção de apneia do sono com análise por-segmento.	23
2.6	Estudos de detecção de apneia do sono com análise por-gravação.	23
3.1	Distribuição dos segmentos por fase do sono.	27
3.2	Atributos calculados do eletrocardiograma.	29
3.3	Exemplo de uma matriz de confusão para um problema binário.	35
3.4	Resultados da classificação de quatro classes.	38
3.5	Resultados detalhados da classificação de quatro classes.	38
3.6	Resultados da classificação de três classes.	39
3.7	Matriz de confusão obtida.	39
3.8	Resultados detalhados da classificação de três classes.	40
3.9	Número e percentagem de segmentos removidos.	40
3.10	Resultados da classificação de três classes com seleção de atributos.	41
3.11	Estudos de classificação do sono com o eletrocardiograma.	43

Acrónimos

ACC	Exatidão (<i>Accuracy</i>)
AED	Atividade Eletrodérmica
ANN	Rede neuronal artificial (<i>Artificial Neural Network</i>)
API	<i>Application Programming Interface</i>
AUC	Área sob a curva ROC (<i>Area Under the ROC Curve</i>)
BLE	<i>Bluetooth Low Energy</i>
BLSTM	Memória bidirecional de longo prazo (<i>Bidirectional Long Short-Term Memory</i>)
CNN	Rede neuronal convolucional (<i>Convolutional Neural Network</i>)
DFA	<i>Detrended Fluctuation Analysis</i>
DNN	Rede neuronal profunda (<i>Deep Neural Network</i>)
ECG	Eletrocardiograma
EEG	Eletroencefalograma
EMG	Eletromiograma
EOG	Eletro-oculograma
FC	Frequência Cardíaca
FIR	Resposta de impulso finita (<i>Finite Impulse Response</i>)
HF	Alta frequência (<i>High Frequency</i>)
HMM	Modelo oculto de Markov (<i>Hidden Markov Model</i>)
IAH	Índice Apneia-Hipopneia
IBI	Intervalo entre batimentos (<i>Interbeat Interval</i>)
IIR	Resposta de impulso infinita (<i>Infinite Impulse Response</i>)
KNN	<i>K-Nearest Neighbors</i>
KSS	Escala de sonolência de Karolinska (<i>Karolinska Sleepiness Scale</i>)
LDA	Análise discriminante linear (<i>Linear Discriminant Analysis</i>)
LF	Baixa frequência (<i>Low Frequency</i>)
LOSO	<i>Leave-One-Subject-Out</i>
LSTM	Memória de longo prazo (<i>Long Short-Term Memory</i>)
NREM	Movimento não rápido dos olhos (<i>Non-Rapid Eye Movement</i>)
PPG	Fotopletismografia (<i>Photoplethysmograph</i>)
PREC	Precisão
PSD	Densidade espectral de potência (<i>Power Spectral Density</i>)
QDA	Análise discriminante quadrática (<i>Quadratic Discriminant Analysis</i>)
REM	Movimento rápido dos olhos (<i>Rapid Eye Movement</i>)
RF	<i>Random Forest</i>
RNN	Rede neuronal recorrente (<i>Recurrent Neural Network</i>)
ROC	Característica de operação do receptor (<i>Receiver Operating Characteristic</i>)
ROS	<i>Random Over-Sampling</i>
SAOS	Síndrome de Apneia Obstrutiva do Sono

SDK	<i>Software Development Kit</i>
SENS	Sensibilidade
SMOTE	<i>Synthetic Minority Over-sampling Technique</i>
SNA	Sistema Nervoso Autónomo
SNP	Sistema Nervoso Parassimpático
SNS	Sistema Nervoso Simpático
SPEC	Especificidade (<i>Specificity</i>)
SVM	<i>Support Vector Machine</i>
UDT	Teste dependente do utilizador (<i>User-Dependent Test</i>)
UIT	Teste independente do utilizador (<i>User-Independent Test</i>)
VC	Validação Cruzada
VFC	Variabilidade da Frequência Cardíaca

Capítulo 1

Introdução

1.1 Contexto

A condução é uma atividade complexa que exige consideráveis capacidades perceptivas, físicas e cognitivas do condutor [1]. É considerado um processo dinâmico em que os principais elementos são o condutor, o veículo e o ambiente de condução. Assim, a função principal do condutor consiste em permanecer consciente do ambiente, tomar decisões e executar ações [2]. A sua atenção ativa desempenha um papel fundamental na condução segura. Por conseguinte, o desenvolvimento de monitorização contínua tem um impacto direto no número de acidentes rodoviários.

As colisões de veículos automotores causam mais de 1.2 milhão de mortes em todo o mundo e um número ainda maior de lesões não fatais a cada ano, afetando negativamente a saúde e o bem-estar dos sobreviventes e suas famílias [3]. As causas mais comuns para acidentes de viação podem ser sumariadas em distração, fadiga e estilo de condução agressivo [2]. Os condutores em estado de sonolência normalmente apresentam decréscimos de vigilância e desempenho. Representando 20% de todos os acidentes rodoviários globais, a sonolência na condução é uma das principais causas de acidentes mortais [4].

Vários fatores podem contribuir para a sonolência ao volante. As suas causas mais frequentes incluem distúrbios do sono, nomeadamente a apneia do sono, ou fatores comportamentais, como privação do sono e trabalho por turnos [5]. Longas horas de condução e a hora do dia são identificados como principais fatores, e em conjunto com a restrição de sono, constituem potenciais fontes de medidas corretivas destinadas a evitar tais ocorrências [6].

É de notar que os condutores profissionais são um dos grupos de risco para estes acidentes. Os resultados de diferentes pesquisas em todo o mundo demonstram que mais de 50% dos motoristas de longo curso já adormeceram ao volante em algum momento [7]. Num estudo acerca dos motoristas de camiões em Portugal, mais de 8 em cada 10 condutores relataram conduzir com sono [8]. A condução é exercida durante longas horas, por vezes irregulares e em conflito com o ritmo circadiano natural do motorista, e associada a estradas longas, monótonas e de alta velocidade. Assim, estes condutores podem acumular uma dívida considerável de sono [9]. O horário de trabalho traz também desafios na manutenção de um estilo de vida saudável, incluindo uma dieta nutritiva e

exercício regular, o que influencia substancialmente o seu bem-estar e desempenho no trabalho [10]. Estudos demonstram que nestes condutores as prevalências de apneia de sono obstrutiva são geralmente mais elevadas do que na população em geral, e apresentam assim, um maior risco de acidentes por sonolência [11]. Por fim, no caso de transporte de passageiros, colocam em perigo um grande número de pessoas, tendo assim um maior impacto a nível da saúde pública.

Em conclusão, a influência da sonolência ao volante em acidentes rodoviários é de gravidade reconhecida e, portanto, um problema que precisa de ser solucionado.

1.2 Motivação

O desenvolvimento de tecnologias para detetar ou prevenir a sonolência ao volante apresenta-se como um grande desafio no domínio dos sistemas de prevenção de acidentes. A monitorização e assistência à condução têm sido progressivamente integradas nos veículos de forma a auxiliar os condutores para uma condução confortável e segura [12]. No entanto, a maioria das abordagens atuais está focada na deteção de um estado comprometido do condutor, e não na sua previsão [13]. Assim, é importante realçar a distinção entre estes dois termos. O objetivo ideal deve ser a previsão do início da sonolência, pois no momento de deteção, a condução sonolenta já pode ter levado a uma situação potencialmente perigosa ou até a um acidente [14].

A privação crónica do sono tornou-se num problema de saúde relevante na sociedade moderna, com inúmeras consequências negativas para a saúde [15]. A prevalência de distúrbios do sono tem aumentado recentemente [16], sendo que a apneia de sono obstrutiva afeta entre 9 a 38% da população adulta em geral [17]. Vários métodos para monitorizar o sono e diagnosticar distúrbios do sono foram desenvolvidos, destacando fatores como acessibilidade e baixo custo. Neste contexto, existe um aumento dos dispositivos disponíveis para os utilizadores, como *wearables* e aplicações móveis. Contudo, estes novos sistemas são raramente validados com a polissonografia, realizada em ambiente clínico e considerada o método padrão-ouro para avaliar o sono, de forma a garantir a sua fiabilidade e validade [18, 19].

1.3 Objetivos

Esta dissertação está inserida no projeto “Sono Ao Volante 2.0” com cooperação entre a FEUP, o Instituto Politécnico do Cávado e do Ave, a empresa Optimizer e o Instituto de Sono. Este aborda o problema da sonolência na condução, e pretende desenvolver uma solução focada na prevenção e segurança rodoviária. Assim, têm como objetivo um sistema que permita a previsão da sonolência ao volante, e a deteção de privação crónica do sono e distúrbios do sono, em particular, a apneia do sono e perturbações do ritmo circadiano. A sua operacionalização será através da análise de dados biométricos e físicos recolhidos por um dispositivo de pulso usado em permanência, nomeadamente, o ritmo cardíaco, pressão de oxigénio arterial e movimento. Deste modo, são incluídos os conceitos de baixo custo e utilização não intrusiva no quotidiano das pessoas. Este sistema é

ilustrado na figura 1.1. É de referir ainda a importância dos dados classificados e validados por profissionais de saúde, essenciais para o cumprimento dos objetivos, que incluem:

- Construir padrões relacionados com os estados de vigília e sono de modo a sinalizar ciclicamente potenciais distúrbios de sono ou privação crónica do sono.
- Monitorizar o estado de vigília durante a condução, e emitir um alerta quando existe previsão de que o estado de sono ocorrerá num curto prazo.

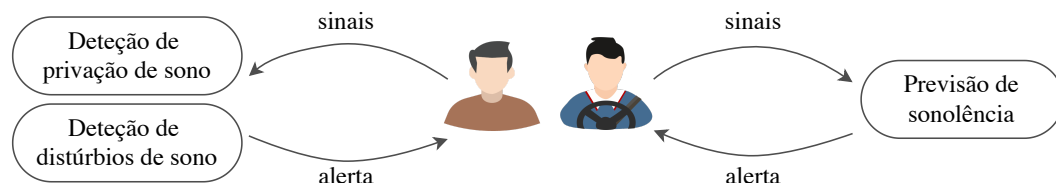


Figura 1.1: Arquitetura geral da solução.

1.4 Contribuição

No âmbito desta dissertação, através de uma análise detalhada da literatura, são apresentados de forma clara conceitos associados ao sono e distúrbios de sono, e realizada uma revisão das soluções atuais, tanto científicas como comerciais, para lidar com as questões abordadas no projeto. Além disso, como trabalho inicial, foi considerada a análise do eletrocardiograma para a classificação das fases do sono. Para tal, foi utilizado um *dataset* público, e comparado o desempenho obtido com diferentes estratégias.

1.5 Estrutura do documento

Além da introdução, este documento contém 3 capítulos, e está organizado da seguinte forma: no capítulo 2 é apresentada uma revisão do estado da arte. São descritos os conceitos fundamentais do sono, seguido da análise dos métodos usados na deteção e previsão de sonolência na condução, classificação do sono, e deteção de apneia. Além disso, são estudados dispositivos e *datasets* para este trabalho. O capítulo 3 relata o trabalho desenvolvido. Este inclui a descrição da metodologia usada e a discussão dos resultados obtidos. Por fim, no capítulo 4, são apresentadas conclusões e sugestões de trabalho futuro.

Capítulo 2

Revisão da literatura

2.1 Sono e ritmo circadiano

O sono é um estado de repouso fundamental para o ser humano, uma necessidade fisiológica essencial que deve ser satisfeita de forma a garantir a sua sobrevivência [20]. É indispensável para o desempenho de atividades básicas da vida diária como a memorização, aprendizagem, produtividade e concentração [21]. Durante o sono ocorrem processos comportamentais, fisiológicos e neurocognitivos, considerados complexos, ativos e altamente organizados, suscetíveis de serem prejudicados pela sua ausência [20, 22].

A quantidade e qualidade do sono têm um grande efeito na saúde e bem-estar das pessoas. A sua restrição, particularmente quando crónica, está associada a défices neurocomportamentais, significativos e cumulativos, como lapsos de atenção, alterações da memória de trabalho, alterações de humor e funções cognitivas reduzidas [23].

A privação parcial do sono pode ocorrer devido à sua fragmentação ou restrição. A fragmentação do sono impede o mesmo de ser fisiologicamente consolidado, interrompendo a progressão normal das suas fases, o que pode ocorrer em certos distúrbios do sono (p.e. apneia do sono obstrutiva). Esta fragmentação pode ser isolada numa fase específica do sono e ser, portanto, referida como privação seletiva do sono. Por outro lado, a restrição é caracterizada pela redução da duração do sono, sendo também referida como dívida de sono. [23]

A privação de sono aumenta os lapsos comportamentais durante tarefas de desempenho cognitivo, que se refletem em episódios de micro-sono [24]. Estes consistem em breves períodos sem resposta, associados a perda de atenção e olhares em branco, sem consciência do indivíduo [25]. A variabilidade do desempenho cognitivo envolve tanto erros de omissão (lapsos evidenciados como falha de resposta oportuna a um estímulo) como de comissão (resposta a um estímulo errado ou inexistente). A sua causa está associada à “instabilidade do estado de vigília”, que ocorre quando os mecanismos homeostáticos de início do sono interferem repetidamente com a manutenção da vigília. Assim, a capacidade do cérebro de manter a atenção é prejudicada pela ativação dos processos do sono. Com o prolongar da restrição de sono, os lapsos podem variar em duração de 0.5

segundos para mais de 10 segundos, aumentando o risco de erros e acidentes em tarefas de vigiância psicomotora, como a condução [24, 23]. A privação de sono pode resultar ainda num número de alterações fisiológicas que contribui para o aumento de riscos para a saúde, nomeadamente a obesidade, diabetes, hipertensão e doenças cardiovasculares [26].

A vulnerabilidade neurocomportamental e fisiológica à restrição do sono está subjacente à variabilidade individual, o que demonstra um padrão de resposta fenotípica [27]. Estes fenótipos observados são dependentes da tarefa, sugerindo que existem perfis cognitivos distintos com vulnerabilidades diferentes no mesmo indivíduo. Da mesma forma, à medida que a restrição de sono continua ao longo do tempo, as diferenças interindividuais no grau de défices cognitivos também aumentam acentuadamente [23].

Estudos sobre a base biológica da vulnerabilidade à privação de sono incidem num modelo de dois processos de regulação do sono, utilizado para descrever os perfis temporais de sono e vigília. O modelo consiste num processo homeostático do sono (S) e num processo circadiano (C), representados na figura 2.1. A sua interação determina o momento de início e duração do sono, assim como a estabilidade das funções neurocognitivas de vigília. O processo S representa a necessidade de dormir, que se desenvolve durante a vigília e dissipa durante o sono. O processo C define a tendência para dormir, representando assim uma modulação oscilatória diária dos limites em que a pressão homeostática aciona o sono [27, 24].

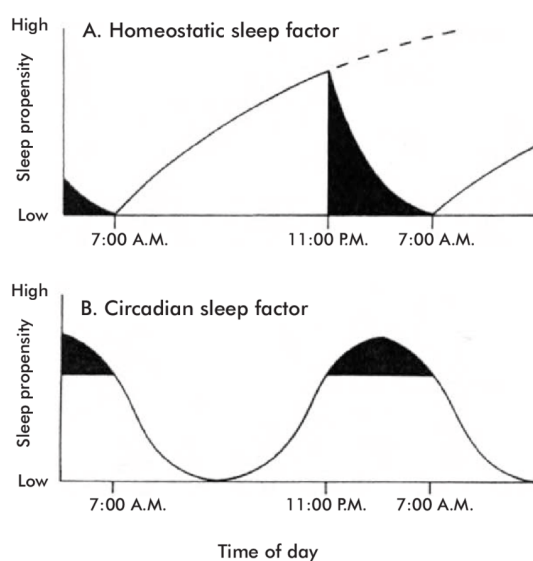


Figura 2.1: Modelo de regulação do sono. Fonte: [28]

À medida que o sono é reduzido, o processo S é acumulado gradualmente ao longo de vários dias. Deste modo, a privação de sono pode elevar a pressão homeostática ao ponto de as funções cognitivas de vigília serem afetadas, mesmo em momentos de pico do ritmo circadiano. Estudos sugerem que os indivíduos mais vulneráveis à restrição de sono experienciam um maior aumento da pressão homeostática ou, alternativamente, apresentam um componente circadiano mais fraco que não a pode neutralizar efetivamente durante a vigília prolongada [27, 24].

O ritmo circadiano é controlado por um relógio biológico endógeno localizado no hipotálamo, sincronizado com o ciclo claro-escuro gerado pelo movimento de rotação da terra [27]. Além do sono, regula processos fisiológicos essenciais como a temperatura corporal, metabolismo, frequência cardíaca, pressão arterial e níveis hormonais [29]. O funcionamento adequado deste sistema de temporização e sincronização permite uma harmonização com os ciclos ambientais [30]. No entanto, estimativas sugerem que mais de 50% da população nas sociedades industrializadas vive permanentemente com um relógio circadiano inadequado ao ambiente externo [31].

A sociedade moderna impõe alguns desafios ao ritmo biológico, nomeadamente a adaptação aos horários sociais mais habituais [30]. Estes interferem consideravelmente nas preferências individuais do sono, denominadas de cronotipos, na maioria da população. A discrepância entre o tempo social e biológico, pode ser descrita como “*jetlag* social”. O início do sono tardio combinado com o despertar cedo, traduz-se no acumular de uma dívida substancial de sono durante a semana de trabalho, compensada depois com o prolongamento da duração do sono. Os cronotipos variam consoante uma predisposição mais matinal (matutino) ou noturna (vespertino). Pesquisas recentes demonstram que estes cronotipos são características herdadas geneticamente e sofrem uma adaptação às condições ambientais. As perturbações do ritmo circadiano surgem de um padrão crónico ou recorrente de distúrbios do sono-vigília devido à disfunção do ritmo circadiano ou desalinhamento entre este e o ambiente externo, causando sintomas como insónia e sonolência excessiva durante o dia [32].

2.2 Padrões de sono

O sono humano é composto por dois estados distintos: movimento não rápido dos olhos (NREM) e movimento rápido dos olhos (REM). Estes estados alternam ciclicamente, sendo que uma noite típica envolve quatro ou cinco ciclos de sono, cada um com duração de cerca de 90 minutos [33]. O hipnograma da figura 2.2 mostra a transição entre as várias fases do sono ao longo de um sono de oito horas.

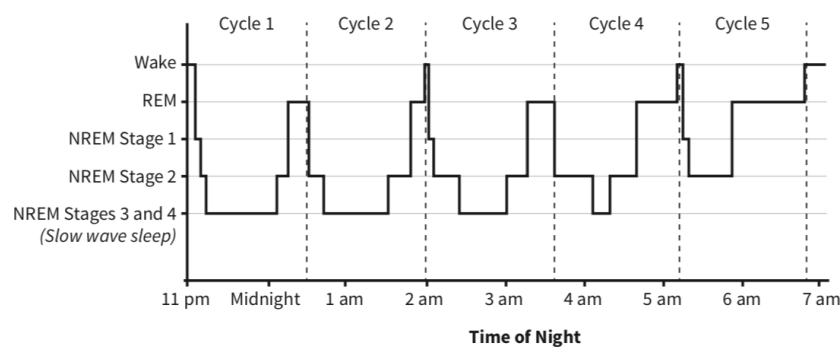


Figura 2.2: Arquitetura do sono. Fonte: [34]

O sono NREM é dividido em três fases, ao longo das quais o sono se aprofunda gradualmente. Quando o sono é iniciado, ocorre a fase um, com a diminuição da maioria das atividades cognitivas e fisiológicas, sendo o estado relacionado com a sonolência. A fase dois e três estão associadas ao sono leve e profundo, respetivamente. À medida que estas fases se sucedem, as ondas cerebrais tornam-se progressivamente mais lentas e o indivíduo fica cada vez menos reativo aos estímulos sensoriais. Conforme se transita novamente para um sono mais leve, o ciclo termina com o sono REM. Durante esta fase, a frequência respiratória e atividade cerebral aumentam, juntamente com a ocorrência de sonhos. [33, 35]

A necessidade diária de sono varia de acordo com a idade e de forma individual [36]. Os ciclos de sono nos adultos apresentam um padrão no qual passam 30% do tempo a sonhar, 20% em sono profundo e 50% em sono leve [35]. O sono é mais profundo na primeira metade da noite, onde existe predomínio da fase NREM, enquanto na segunda metade ocorre predomínio das fases mais superficiais do sono NREM e de sono REM. O sono profundo é considerado restaurador, pois está associado a reduções acentuadas na atividade simpática, resultando numa respiração estável e na diminuição da frequência cardíaca e pressão arterial [20].

A polissonografia é o método mais utilizado para a avaliação objetiva do sono, nomeadamente na identificação de distúrbios de sono [37]. São incluídos, entre outros, os seguintes parâmetros:

- eletroencefalograma (EEG) - atividade elétrica cerebral;
- eletro-oculograma (EOG) - movimentos oculares;
- eletromiograma (EMG) - atividade muscular;
- eletrocardiograma (ECG) - atividade cardíaca;
- fluxo oro-nasal e movimentos torácicos e abdominais - respiração;
- oximetria - oxigenação do sangue.

A figura 2.3 mostra as leituras poligráficas típicas associadas às diferentes fases do sono.

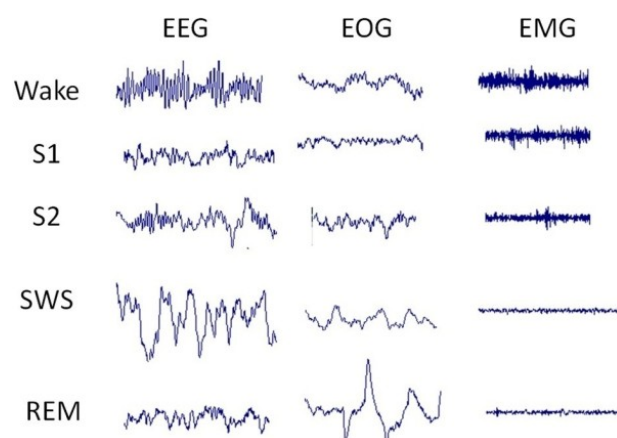


Figura 2.3: Leituras poligráficas das fases do sono. Fonte: [38]

2.3 Métodos para detecção de sonolência

Várias técnicas para estimar a sonolência foram propostas na literatura. Estes métodos podem ser classificados em diferentes categorias, de acordo com a fonte de informação: medidas subjetivas, baseadas nos veículos, comportamentais e fisiológicas [39, 13]. No entanto, uma abordagem híbrida que combina vários métodos pode também ser considerada.

As medidas subjetivas são baseadas na estimativa pessoal do condutor e são geralmente avaliadas recorrendo a questionários. A métrica mais utilizada é a Escala de Sonolência de Karolinska (KSS), representada por uma escala de nove pontos que varia de “extremamente alerta” a “extremamente sonolento” [40], como mostrado na figura 2.4. Durante estes testes, os indivíduos são solicitados a classificar o seu estado de alerta de acordo com a escala indicada. A autoavaliação geralmente é realizada repetidamente durante o tempo de teste, através de um intervalo de tempo ou determinadas condições [41]. Uma limitação destas medidas é a inviabilidade de implementação em condições de condução real, não sendo possível detetar mudanças repentinas na sonolência, por exemplo, quando o KSS é usado a cada 5 minutos [42]. Além do mais, pedir aos condutores que avaliem o seu estado de sonolência atual pode estimular a atenção e, portanto, influenciar as classificações. Alguns estudos demonstram que estas medidas apresentam baixa correlação e dissociação com medidas objetivas [43].

Rating	Verbal descriptions
1	Extremely alert
2	Very alert
3	Alert
4	Fairly alert
5	Neither alert nor sleepy
6	Some signs of sleepiness
7	Sleepy, but no effort to keep alert
8	Sleepy, some effort to keep alert
9	Very sleepy, great effort to keep alert, fighting sleep

Figura 2.4: Escala de Sonolência de Karolinska. Fonte: [40]

As medidas baseadas nos veículos fornecem uma avaliação do desempenho da condução através da análise das capacidades do condutor e da forma como este controla o veículo [43]. O movimento do volante e o desvio padrão da posição da faixa são as métricas mais utilizadas [13]. Estas medidas apresentam a vantagem de serem fiáveis, não intrusivas e em tempo real. Contudo, uma grande limitação é que estas variações ocorrem tipicamente apenas nas fases finais da sonolência, podendo ser tarde demais para impedir a ocorrência de acidentes. Além disso, são extremamente dependentes da geometria da estrada, sendo eficazes num número limitado de condições [43]. Em consequência, geralmente falham quando expostas a ambientes reais com variação substancial. É

de referir ainda que estas medidas não são indicadores específicos de sonolência, e podem variar também com fatores como a distração cognitiva ou visual [44].

As medidas comportamentais são focadas na capacidade de concentração do condutor, detetadas principalmente por técnicas de processamento de imagem [43]. A fonte de informação visual pode incluir a expressão facial, movimento dos olhos e da cabeça, ou o bocejar. Geralmente, estes sistemas usam uma câmara de vídeo para aquisição de imagem e após a extração dos parâmetros de interesse, técnicas de *machine learning* são usadas para determinar o nível de sonolência [45]. A deteção do estado dos olhos tem sido o foco principal da maioria dos estudos publicados [39]. Em particular, alguns métodos medem a frequência do piscar dos olhos, que normalmente é cerca de 10 vezes por minuto, e diminui nos condutores sonolentos [45]. Uma das métricas mais utilizadas é o PERCLOS, definida como a proporção durante um período de tempo especificado (p.e. 3 min) em que os olhos estão fechados 80% ou mais, refletindo assim, o fechar lento dos olhos [46]. Embora não intrusivos, estes métodos apresentam desafios significativos que influenciam a sua precisão e credibilidade [47]. Nomeadamente, as capacidades de deteção são fortemente afetadas por variações nos fatores ambientais e nas condições de condução, como por exemplo, mudanças de iluminação dentro ou fora do veículo e o uso de óculos [43].

A última categoria de medidas envolve o uso de sinais fisiológicos. Estes apresentam alterações nas fases iniciais da sonolência e são, portanto, mais adequados para detetar a sonolência de forma atempada [46]. No entanto, é difícil definir uma relação direta com a sonolência, pois estas medidas variam com outros estados (incluindo emoções, carga de trabalho ou fadiga física), contexto ou de forma individual, pelo que não devem ser considerados de forma isolada [13]. Apesar da maior fiabilidade e precisão comparativamente aos outros métodos, a natureza intrusiva da medição de sinais fisiológicos dificulta a sua utilização em condições reais de condução. Deste modo, o uso de dispositivos sem fios pode ser considerado, contudo, com a desvantagem de uma menor precisão [39].

Por fim, as abordagens híbridas combinam estes métodos de modo a alcançar o melhor desempenho possível. Um resumo das vantagens e limitações de cada método é apresentado a seguir, na tabela 2.1.

Medidas	Vantagens	Limitações
Subjetivas	Considera a opinião pessoal	Inviável em tempo real
Baseadas nos veículos	Tempo real, não intrusivo	Indisponível em condução automática
Comportamentais	Não intrusivo	Precisão interpessoal
Fisiológicas	Fiável	Intrusivo

Tabela 2.1: Comparação entre medidas para deteção de sonolência. Adaptado de [42]

2.4 Análise de sinais fisiológicos

As medidas fisiológicas são consideradas objetivas e frequentemente utilizadas para a detecção de sonolência. Esta detecção representa-se como um problema de classificação, nomeadamente, se o condutor está ou não sonolento. No entanto, a evolução do seu estado operacional também pode ser considerada um problema de regressão (ou seja, o condutor passa por vários estados contínuos), embora estes modelos sejam raramente utilizados na literatura [13]. Assim, a aplicação de sinais fisiológicos em tarefas de classificação é dividida em várias etapas, representadas na figura 2.5 para a identificação de sonolência [48, 2].

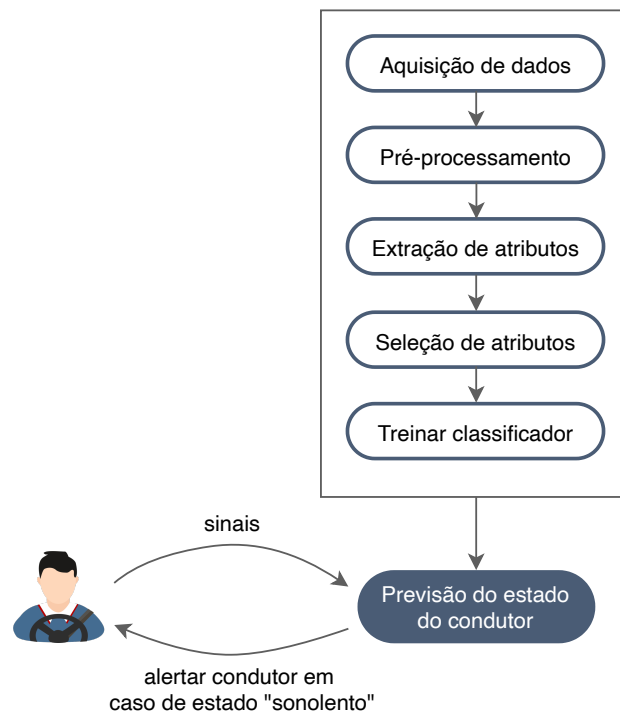


Figura 2.5: Detecção de sonolência no condutor baseada em sinais.

Depois de obtido, o sinal passa por um pré-processamento que inclui a filtragem e rejeição de artefactos, de modo a melhorar a sua qualidade e não perder informação relevante. Esta tarefa é tipicamente realizada através de filtros digitais, que podem ser de dois tipos: resposta de impulso finita (FIR) ou infinita (IIR). O procedimento de extração de atributos (do inglês *features*) é considerado o mais importante. Estes podem ser classificados em quatro grupos:

1. Domínio do tempo: as medidas estatísticas são dos atributos mais simples que podem ser derivados da análise no domínio do tempo (p.e. média, variância e assimetria). Ao aplicar estas medidas, cada intervalo de tempo é considerado um processo univariado que ignora correlações entre intervalos.
2. Domínio da frequência: os atributos no domínio da frequência caracterizam a estrutura espectral do sinal e podem ser obtidos através da densidade espectral de potência (PSD). Esta

é tipicamente estimada com base na transformada de Fourier, e descreve como a potência de um sinal é distribuída nas suas componentes de frequência.

3. Domínio tempo-frequência: a maioria dos sinais fisiológicos são não-estacionários, ou seja, as suas propriedades estatísticas são definidas em função do tempo [49]. A transformada de Wavelet representa ambos os domínios de tempo e frequência, sendo frequentemente usada na análise destes sinais.
4. Não lineares: estes atributos são capazes de capturar fenómenos fisiológicos complexos de forma eficaz, utilizando técnicas não lineares [50].

Caso necessário, métodos de seleção de atributos ou redução de dimensionalidade são usados para reduzir o número de atributos, ou gerar atributos de baixa dimensão. Por fim, são aplicados algoritmos de classificação e o desempenho é avaliado ao comparar cada previsão com o valor de referência (do inglês *ground-truth*). Para tal, são consideradas métricas como a exatidão, sensibilidade e especificidade.

Nas subsecções seguintes são descritos os sinais frequentemente utilizados no estudo do sono.

2.4.1 Eletroencefalograma

O EEG é utilizado para medir a atividade elétrica do córtex cerebral [48]. O sinal obtido inclui as seguintes bandas de frequência:

- delta (0.5–4 Hz): ocorrência do sono;
- teta (4–8 Hz): sonolência;
- alfa (8–13 Hz): relaxamento e criatividade;
- beta (13–30 Hz): estado de alerta.

Este sinal é considerado um dos métodos mais fiáveis para a deteção de sonolência [43], observada através de uma diminuição da frequência alfa e um aumento da frequência teta [46]. Além de diferenciar os estados de vigília e sono, o EEG é também utilizado para identificar as diferentes fases do sono [2].

2.4.2 Eletrocardiograma

O ECG regista a atividade elétrica do coração [39]. Permite obter a informação associada à sonolência e fases do sono, descrita de seguida [2, 46]:

1. Frequência cardíaca (FC): o número de batimentos cardíacos por minuto varia significativamente entre as diferentes fases do sono, de acordo com o equilíbrio entre os dois componentes do sistema nervoso autónomo (SNA): o sistema nervoso simpático (SNS) e parassimpático (SNP). Os estados de vigília são caracterizados por um aumento da atividade simpática ou diminuição da atividade parassimpática, e o oposto para estados de relaxamento [51]. Ao transitar de um estado alerta para um estado de sono, a FC geralmente diminui.

2. Variabilidade da frequência cardíaca (VFC): representa a alteração no intervalo de tempo entre dois batimentos cardíacos sucessivos. Este é referido como intervalo RR, e ilustrado na figura 2.7.

A atividade do SNA pode ser efetivamente detetada através da análise da VFC. Tipicamente, em atividades consideradas mentalmente difíceis e fisicamente fáceis, a FC aumenta e a VFC diminui. A partição do espectro de potência da VFC típico é mostrada na figura 2.6.

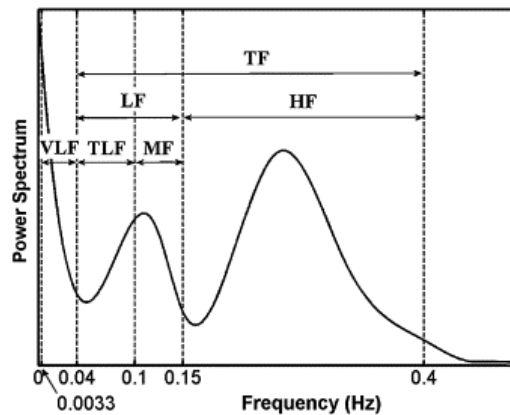


Figura 2.6: Partição do espectro de potência da variabilidade da frequência cardíaca. Fonte: [52]

A sua análise permite identificar a sonolência de forma antecipada, utilizando as potências das bandas de baixa e alta frequência (LF e HF) [53]:

- LF (0.04–0.15 Hz): reflete a atividade do SNS e SNP.
- HF (0.15–0.4 Hz): reflete a atividade do SNP.

Desta forma, a relação LF / HF expressa o equilíbrio entre a atividade do SNS e SNP, e diminui à medida que um indivíduo progride para um estado sonolento. Esta diminuição é também verificada durante o sono NREM, com a redução gradual da FC e VFC [52]. No sono REM é observado o oposto.

3. Frequência respiratória: a atividade respiratória pode ser estimada a partir da banda HF da VFC. Nos indivíduos sonolentos, a frequência respiratória tende a diminuir, enquanto que a regularidade tende a aumentar [54]. Por outro lado, no sono REM a frequência da respiração é irregular, e um padrão rápido e superficial tende a prevalecer [55].

2.4.3 Eletro-oculograma

O EOG mede a diferença de potencial elétrico entre a córnea e a retina do olho humano, alterada pela atividade ocular. Deste modo, a sonolência pode ser detetada através de indicadores baseados no movimento das pálpebras (p.e. amplitude, duração e frequência do piscar) ou dos olhos (p.e. movimento lento e rápido dos olhos). [2]

2.4.4 Eletromiograma

O EMG avalia o sinal elétrico gerado pela contração muscular. Estudos revelam que a amplitude dos sinais EMG diminui gradualmente com a fadiga muscular, e que esta está correlacionada com a sonolência. A principal desvantagem deste sinal consiste na sua natureza aleatória e complexa e na dependência das propriedades biológicas e estruturais do músculo. [2]

2.4.5 Fotopletismografia

A fotopletismografia (PPG) mede a variação volumétrica da circulação sanguínea periférica [56]. O sinal é extraído por um oxímetro de pulso através de luz infravermelha, e permite estimar a saturação de oxigênio no sangue (SpO_2) e a FC. Este método constitui uma alternativa não invasiva ao ECG pois permite medir a última sem o uso de elétrodos. Neste caso, para a análise da VFC, em vez dos intervalos RR são medidos os intervalos entre batimentos (IBI), como demonstrado de seguida, na figura 2.7.

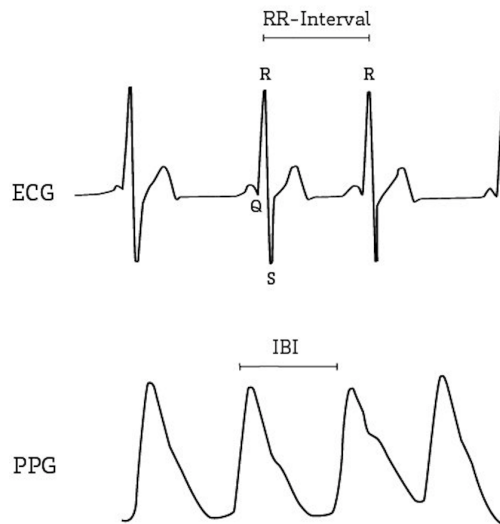


Figura 2.7: Análise de batimentos cardíacos com sinais de eletrocardiograma e fotopletismografia. Fonte: [57]

2.4.6 Propriedades da pele

A atividade eletrodérmica (AED) constitui uma medida da condutividade da pele. Durante a sonolência, é desencadeada a atividade do SNP, o que causa uma redução da transpiração. Consequentemente, a resistividade da pele aumenta e a condutividade diminui. No entanto, esta técnica é altamente sensível à humidade e temperatura atmosféricas. [2]

A temperatura da pele é controlada pelo sistema de termorregulação do corpo humano e varia com o nível de sonolência. Estudos demonstram que um aumento inicial da temperatura no pulso seguido de uma diminuição gradual é observado nos condutores sonolentos [58]. Além disso, no sono NREM a temperatura da pele aumenta, e a temperatura interna diminui [59].

2.5 Detecção de sonolência

Atualmente estão disponíveis vários sensores para monitorização da saúde de forma não invasiva e em tempo real. Assim, a tendência dos dispositivos usáveis (do inglês *wearable*) prevalece e expande-se continuamente, sendo estimada o seu uso por uma em cada seis pessoas [60]. A ampla aceitação é aliada a vantagens como o baixo custo, facilidade de utilização e medição de múltiplos sinais fisiológicos. Existe, portanto, uma motivação para analisar a viabilidade dos dispositivos de pulso como os relógios ou pulseiras na deteção da sonolência na condução. Deste modo, o seu âmbito poderia ser aumentado além da monitorização diária da saúde ou condicionamento físico, constituindo uma solução sem necessidade de instalar sensores adicionais nos veículos. Assim, os estudos recentes [60] e [61] avaliam a aplicabilidade destes dispositivos na classificação binária da sonolência. Para tal, é considerada a análise da VFC obtida de um sensor PPG.

Em [60], o estado do condutor foi avaliado em ambiente de condução simulado utilizando o KSS, no qual os níveis 1–4 foram considerados como alerta e os níveis 7–9 como sonolência. O sinal PPG foi analisado com segmentos de 2 min e incremento de 2 s, e a fiabilidade foi confirmada pela alta exatidão dos modelos obtidos, atingindo 99.9%. Além disso, foi possível observar que os padrões de sonolência variam de acordo com a faixa etária assim como a VFC no contexto da deteção de sonolência. Nomeadamente, os condutores jovens são mais propensos à sonolência e os modelos treinados com apenas dados de jovens não são suficientes para serem aplicados em condutores de mais idade. Assim, é sugerido o desenvolvimento de modelos gerais, considerando dados de diferentes faixas etárias, combinados com modelos individuais.

Em [61], para avaliar o estado do condutor, foi aplicado um processo de duas fases através da combinação da classificação de vídeo e processamento de imagem. Em cada intervalo de 5 min, 1 min foi extraído de forma aleatória com o objetivo de eliminar a dependência de tempo entre segmentos individuais, e a classificação de sonolência foi realizada numa escala de 1–6 por dois observadores treinados. De modo a identificar possíveis ocorrências de micro-sono nos restantes 4 min, foram detetados eventos associados ao fechar lento dos olhos, e atualizadas as classificações iniciais. A classe de alerta agrupou os níveis 1–3, e a classe de sonolência os níveis 4–6. Para a análise da VFC, o sinal PPG foi processado no domínio do tempo, frequência e não linear, em segmentos de 5 min com um incremento de 2 s. Tendo em conta que a capacidade de generalização é uma questão crucial devido à variação dos sinais fisiológicos entre condutores, foram usados testes dependentes e independentes do indivíduo, nomeadamente, validação cruzada (VC) estratificada *10-fold* e *leave-one-subject-out* (LOSO). Foram aplicados diversos classificadores, e os resultados fornecem evidências de que a configuração não intrusiva alcança uma exatidão semelhante à dos dispositivos de nível médico, ou seja, ao ECG. Concretamente, estes testes permitiram atingir uma exatidão de 92% e 73%, pela mesma ordem.

Em conclusão, observa-se que a sonolência pode ser efetivamente detetada através da análise da VFC. Alguns estudos consideram outras medidas, como apresentado na tabela 2.2. No entanto, com foco na condução automática futura, sensores como o acelerómetro ou giroscópio serão mais difíceis de usar, uma vez que padrões de movimento não podem ser avaliados continuamente.

Ref.	Ano	Medidas	Atributos	Segmentos	Valor de referência	Algoritmo	No. classes	Validação	Exatidão
[67]	2014	Movimento (acelerómetro, giroscópio)	Aceleração linear, velocidade angular	1 min	-	Baseado em <i>threshold</i> (nº de ajustes por min determinados por SVM binário)	-	Identificação de sonolência pelo PERCLOS	98.8%
[68]	2015	Movimento, PPG, AED	5 atributos incluindo FC, VFC e nível de stress	1 min com incremento de 1 s	KSS observada a cada 1 min	SVM	5	-	98.3%
[69]	2016	Movimento	8 atributos no domínio do tempo frequência e fase	1 min com incremento de 1 s	KSS observada a cada 1 min	SVM	5	VC <i>holdout</i> (70-30)	98.15%
[70]	2018	Movimento, PPG, AED, temperatura	38 atributos com seleção	10 s com sobreposição de 2 s	Escala 1-5 com base na análise de expressões faciais a cada 5 min	SVM	4 (normal, sonolência, stress, fadiga)	VC <i>5-fold</i> , VC <i>LOSO</i>	98.34% 68.31%
[71]	2019	PPG	3 atributos associados à VFC	2 min	Análise do rosto e comportamento a cada 1 min	CNN	2	VC <i>10-fold</i>	64%
[72]	2019	PPG, AED	2 atributos da VFC e AED	40 s	-	Baseado em <i>threshold</i> (sonolência numa escala 1-4 em 4 momentos do dia)	-	Identificação de sonolência pelo condutor	80%

Tabela 2.2: Estudos de deteção de sonolência no condutor. SVM—*support vector machine*; CNN—rede neuronal convolucional.

2.5.1 Soluções comerciais

Além da pesquisa, existem no mercado alguns dispositivos para a detecção de sonolência. O Anti-Sleep Alarm [62] é um anel que avalia os níveis de alerta e concentração do condutor através de sensores AED, e contém alarmes integrados. O Sleep Alert [63] monitoriza os sinais de EEG e EMG, posição da cabeça e do corpo através de um chapéu e pulseira que emitem som e vibração, respetivamente, ao detetar a fadiga. A banda Life [64] pode ser colocada num capacete de modo a avaliar o EEG, e a monitorização é realizada por uma aplicação móvel. O Vigo [65] é um *headset* que analisa mais de 20 parâmetros na variação do piscar dos olhos, e contém um acelerómetro e giroscópio para medir os movimentos da cabeça. Ao detetar a sonolência, estimula a atenção do condutor com música, luz ou chamadas telefónicas. O sistema Eagle Industrial [66] é emparelhado com uns óculos que medem o movimento das pálpebras, e exibe o índice de sonolência num *tablet*. A monitorização em tempo real é também disponibilizada, o que permite aos supervisores observar a pontuação de cada condutor.

Como descrito, não existem dispositivos de pulso com análise da frequência cardíaca. Além disso, a falta de validação das soluções propostas impede a compreensão do seu funcionamento e confiança nos resultados obtidos.

2.6 Previsão da sonolência

O estudo em [13] tem como objetivo determinar se a informação utilizada para detetar a sonolência também a permite prever. Para tal, foram medidos indicadores fisiológicos, comportamentais e baseados no veículo. Diferentes combinações destas informações foram testadas. Através da análise de vídeo, dois avaliadores treinados classificaram cada segmento de 1 min numa escala que varia de 0 a 4, e a sua média foi considerada o nível de sonolência. Foram desenvolvidos dois modelos independentes utilizando redes neuronais. A cada minuto, o modelo de detecção identifica o nível de sonolência e o modelo de previsão indica o tempo necessário para atingir um determinado nível de sonolência (definido em 1.5). O melhor desempenho em ambas as tarefas foi obtido com indicadores comportamentais e informação adicional, nomeadamente, tempo de condução e informação dos participantes. Estes modelos conseguiram detetar e prever a sonolência com um erro quadrático médio de 0.22 e 4.18 min, respetivamente. Além disso, na previsão, o valor absoluto do erro é inferior a 2 min em 95% dos dados de teste e na detecção, inferior a 0.5 em 80% dos casos. Assim, é concluído que, em ambiente controlado e propício à sonolência, é possível prever a dinâmica do estado do condutor. Contudo, a variabilidade interindividual foi apenas considerada em [73]. De forma a encontrar um compromisso entre modelos generalizados e individuais, foi usada aprendizagem adaptativa. Ambos os modelos foram treinados com um grupo de condutores e posteriormente adaptados para cada indivíduo específico. A melhoria no desempenho é significativa a partir dos primeiros 3 min de condução, até um total de 15 min, onde esta atinge cerca de 40% na detecção e 80% na previsão. No entanto, a questão da variabilidade intraindividual não foi abordada, ou seja, com que regularidade esta adaptação seria necessária.

De modo a efetuar a previsão, foi considerada a diferença entre o tempo restante até o nível-alvo ser atingido. No entanto, alguns estudos consideram outras abordagens. Em [74], indicadores subjetivos e fisiológicos foram utilizados na previsão de sonolência durante uma tarefa de vigilância. Para tal, o valor de referência foi deslocado num segmento, e os resultados demonstraram uma correlação significativa. O estudo em [75] analisou a eficácia das medidas fisiológicas e comportamentais para prever um episódio de sonolência com 20 s de antecedência. Com base na avaliação subjetiva do condutor (escala 1–3) a cada 1 min, foi aplicado um modelo de regressão logística multinomial. A maior exatidão foi obtida com o intervalo entre 20 s e 120 s antes da previsão. Esta abordagem permitiu também concluir que a estrutura ótima deste modelo é diferente entre os participantes. Em [76], o micro-sono é detetado com 96% de exatidão e é possível prever a ocorrência do próximo episódio com antecedência de 5 s a 5 min. Para tal, é avaliada a medida $pNN50$ obtida do ECG. No entanto, a principal limitação deste trabalho é que o primeiro micro-sono não pode ser previsto. Para o mesmo objetivo, em [77], a deteção de micro-sono utilizando regressão logística atinge uma exatidão de 93%, considerando o fator individual do condutor e medidas oculares. Foi possível atingir uma especificidade de 98% e sensibilidade de 67%, o que indica poucos falsos alarmes, mas com probabilidade de estes episódios não serem detetados. Além disso, não houve alterações significativas no desempenho do modelo usando diferentes intervalos de tempo de 1 min em relação aos eventos (de 1 min a 10 min antes). No estudo em [78], foi desenvolvido um modelo de tempo de falha acelerado que permite prever o tempo de condução contínua antes do início da sonolência. Para tal, foram considerados fatores ambientais e demográficos como hora do dia, temperatura, experiência de condução, idade e hábitos de sono. Através da sua análise, é possível compreender como a sonolência é influenciada por estes fatores.

Em conclusão, a modelação da sonolência como um processo contínuo permite gerar sistemas de deteção precisos e capazes de prever a sonolência. Contudo, é necessário considerar a variabilidade individual do condutor. Nestes estudos, a maioria das medidas fisiológicas utilizadas foram intrusivas, pelo que se torna interessante averiguar a possibilidade de obter resultados semelhantes de forma simples e não invasiva.

2.7 Classificação do sono

A classificação das fases do sono é essencial para o seu estudo [48]. Este processo é conhecido como estadiamento do sono, e é feito com base no exame de polissonografia, na qual os sinais são analisados em intervalos de 30 s. Deste modo, torna-se um procedimento dispendioso, propenso a erros e demorado. Além disso, como é realizado em ambiente hospitalar, pode afetar a eficiência do sono dos pacientes. Portanto, comprova-se a relevância do desenvolvimento de dispositivos de baixo custo que possam detetar e analisar automaticamente padrões de sono, de forma a identificar condições como os distúrbios do sono.

A crescente consciência da importância do sono levou à maior utilização de monitores de sono, que permitem obter informação útil sobre a sua qualidade. Para a aquisição dos dados, pode ou não ser necessário contacto com o indivíduo, nomeadamente ao optar por dispositivos usáveis ou

aplicações móveis. Contudo, devido a uma maior proximidade dos sensores, os primeiros são considerados mais precisos [79]. Para avaliar a qualidade do sono a longo prazo, é tradicionalmente utilizada a actigrafia [80]. Este método conta com o uso de um dispositivo de pulso, e infere a vigília e o sono através de um acelerómetro ao detetar a presença ou ausência de movimento. No entanto, como não permite descrever a arquitetura do sono, estes dados são tipicamente usados em combinação com a PPG nos dispositivos atuais, que consideram aquisição de dados multisensores.

Embora cada vez mais comuns, estudos demonstram que estes dispositivos tendem a subestimar as interrupções do sono e a sobrestimar o tempo total e eficiência do mesmo, ou seja, priorizam a sensibilidade à especificidade [79]. Em particular, estas medidas encontram-se nos intervalos 95–97% e 39–62%, respetivamente, em quatro soluções comerciais analisadas. Além disso, os algoritmos que usam não são de código aberto, e os dados por tratar obtidos dos sensores raramente são acessíveis para uso externo [80]. Como resultado, a sua utilidade para aplicações clínicas e de pesquisa, apesar de promissora, é limitada [81].

Na tabela 2.3 são sumariados estudos que usam dispositivos de pulso na classificação das fases do sono, considerando modelos independentes do indivíduo. Entre as medidas de desempenho, é calculada a estatística Kappa (κ), que considera o potencial do acaso na concordância entre o valor obtido e esperado. A análise pode ser abordada nas seguintes resoluções:

- 2 classes: vigília (W), sono;
- 3 classes: vigília, NREM (N), REM (R);
- 4 classes: vigília, N1+2 (L), N3 (D), REM;
- 5 classes: vigília, N1, N2, N3, REM.

No estudo em [82], foi possível identificar o sono leve, profundo e REM com sensibilidades de 69%, 62% e 72%, respetivamente. Na classificação das cinco fases, as abordagens [83] e [84] treinaram uma rede neuronal recorrente (RNN) com memória bidirecional de longo prazo (BLSTM). Como o sono apresenta padrões temporalmente sequenciais, ao considerar as suas dependências obtêm um melhor desempenho. Particularmente em [84], exceto na fase N1, as áreas sob a curva (AUC) excedem 0.8. Por último, em [85], foi possível concluir que o PPG introduz uma melhoria na deteção sono-vigília em relação à actigrafia.

Um desafio constante nestes estudos é que abordam um problema de dados não balanceados. Nomeadamente, se os indivíduos do conjunto de treino demonstram uma eficiência de sono alta, pode ser gerada uma especificidade baixa. Por outras palavras, se existirem poucos segmentos de vigília, o classificador tem oportunidade reduzida de os distinguir corretamente e é mais suscetível a ruído. Além disto, exceto no estudo [83], o teste foi limitado a *datasets* de adultos saudáveis, sem a validação adicional para outras faixas etárias ou patologias do sono.

Por fim, alguns estudos consideram métodos alternativos para medir objetivamente o sono, nomeadamente, com sinais EEG, ECG, EMG, AED ou de respiração [86]. O primeiro apresenta o desempenho mais elevado entre os métodos que usam apenas um sinal de canal único. Contudo, a maioria das soluções incorpora medidas intrusivas, o que limita a sua utilização. Em contrapartida, os dispositivos de pulso permitem uma análise contínua e integrada.

Ref.	Ano	Medidas	Atributos	Algoritmo	Validação	2 classes	3 classes	4 classes	5 classes
[85]	2017	ACE, PPG	VFC no domínio do tempo e frequência, e movimento	LDA	VC holdout (2 sets)	Acc: 0.92 Sp: 0.58 Se: 0.97 κ : 0.55	Acc: 0.73 κ : 0.46	Acc: 0.59 κ : 0.42	-
[82]	2017	ACE, PPG	54 atributos da VFC, movimento e respiração	LDA	VC LOSO	-	-	Acc: 0.69 Sp: 0.69 Se: 0.95 κ : 0.52	-
[84]	2018	ACE, PPG	Multi-level feature learning (bag-of-words e k-means)	BLSTM	VC LOSO	-	-	-	Acc: 0.68 F-score: 0.64
[83]	2018	PPG	135 atributos VFC	BLSTM + transfer learning de ECG	VC 4-fold	-	-	-	Acc: 0.75 κ : 0.63
[80]	2019	ACE, PPG, no. passos Apple Watch	FC, movimento e ritmo circadiano	ANN	VC holdout repetida (70-30)	Acc: 0.90 Sp: 0.6 Se: 0.93 κ : 0.45	Acc: 0.72 Sp: 0.60 κ : 0.28	-	-
[88]	2019	PPG	VFC no domínio não linear	SVM + filtro Bayesiano	VC LOSO	-	Acc: 0.73 κ : 0.43	Acc: 0.60 κ : 0.38	Acc: 0.54 κ : 0.35
[89]	2019	ACE, PPG	VFC, movimento, respiração e estatísticas do dataset	LDA	VC LOSO	-	Acc: 0.85 κ : 0.67	Acc: 0.77 κ : 0.58	-

Tabela 2.3: Estudos de classificação do sono. Acc—exatidão; Se—sensibilidade do sono; Sp—especificidade da vigília; ACE—aceleração; LDA—análise discriminante linear; ANN—rede neuronal.

2.7.1 Soluções comerciais

Os monitores de sono permitem medir vários parâmetros do sono como tempo total, qualidade e métricas de saúde que o possam afetar [87]. Alguns exemplos são apresentados na tabela 2.4. Apesar dos dispositivos semelhantes ao EEG obterem um melhor desempenho, é improvável que, no futuro próximo, tenham o mesmo impacto que outros devido a custos relativamente mais altos, uso limitado (não podem ser facilmente usados 24-h por dia) e desafios na gravação de sinais de boa qualidade em condições não controladas e não-laboratoriais [81]. Deste modo, os dispositivos de pulso constituem uma solução promissora.

Tipo	Medidas	Exemplos
Relógio ou pulseira desportiva	FC, movimento	Fitbit [90], Polar [91], Apple Watch [92]
Aplicação móvel	Movimento, som	Sleep Cycle [93], SleepScore [94]
Colocado no colchão	FC, som, movimento	Beddit [95], Withings Sleep [96], Tomorrow Sleeptracker Monitor [97]
Faixa para a cabeça (<i>headband</i>)	EEG, FC, movimento	Dreem [98], Smartsleep [99], Sleep Shepherd [100]
Colocado na mesa de cabeceira	Temperatura, luz, som movimento	SleepScore Max Tracker [101], Resmed S+ [102]

Tabela 2.4: Exemplos de monitores de sono.

2.8 Detecção de apneia do sono

A apneia do sono é um distúrbio respiratório comum, caracterizado por episódios recorrentes de obstrução parcial ou completa das vias aéreas superiores durante 10 s ou mais, que resulta na fragmentação do sono e redução da saturação de oxigénio no sangue [103, 104]. As diferentes condições são descritas de seguida:

- Apneia obstrutiva (SAOS): interrupção da respiração causada pela limitação do fluxo de ar;
- Apneia central: ausência de esforço respiratório;
- Apneia mista: combinação de apneia obstrutiva e central;
- Hipopneia: redução significativa do fluxo de ar, devido a uma respiração lenta e superficial.

O diagnóstico de SAOS é baseado no índice apneia-hipopneia (IAH), calculado pelo número de eventos de apneia e hipopneia por hora durante o sono. Os indivíduos com um IAH acima de 5, combinado com outros sintomas clínicos, como a sonolência diurna excessiva e baixa capacidade cognitiva, são geralmente diagnosticados com SAOS [105]. Além disso, esta pode ser considerada leve (IAH 5-15), moderada (IAH 15-30) ou grave (IAH >30).

Como alternativa à polissonografia, algumas abordagens consideram a análise de um conjunto reduzido de sinais, incluindo o ECG, EMG, EEG, saturação de oxigênio e som respiratório [105]. O primeiro tem sido amplamente utilizado pois é um sinal fisiologicamente relevante, e pode ser facilmente obtido por um dispositivo usável [106]. Assim, os atributos mais comuns são gerados a partir do ECG e da respiração derivada do ECG (EDR) [107]. Em particular, durante um evento de SAOS ocorre uma mudança de energia do sinal em direção a frequências baixas para duas séries temporais distintas: intervalo RR e amplitude RS, ilustradas na figura 2.8.

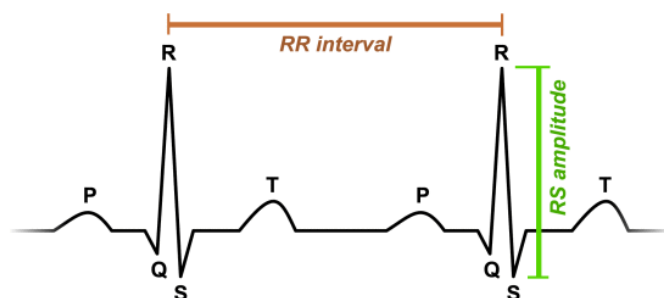


Figura 2.8: Morfologia do eletrocardiograma com o intervalo RR e amplitude RS. Fonte: [107]

Alguns estudos de detecção de SAOS baseada em ECG utilizam o *dataset* “Apnea-ECG” [108]. Os eventos de apneia são identificados com base na análise de segmentos de 1 min. Adicionalmente, a gravação completa de cada indivíduo é classificada de acordo com o IAH. Estas abordagens são apresentadas nas tabelas 2.5 e 2.6. No entanto, a primeira não permite a comparação direta dos resultados devido ao uso de diferentes subconjuntos. Por outro lado, a classificação por gravação considera a mesma amostra, e por isso, o valor de correlação entre o IAH real e obtido é calculado de modo garantir a fiabilidade da comparação.

Estes estudos são focados em diferentes aspectos. Em [107] e [109] é tido em conta o custo e tempo computacional associados a um número elevado de atributos ou análise complexa. Particularmente, em [109] o sinal EDR não é analisado, tornando a abordagem prática para aplicações reais devido à tolerância da distorção do sinal ECG. Outra consideração relevante é o desenvolvimento de algoritmos menos dependentes de conhecimento prévio do domínio. Assim, em [106] e [110], o passo de engenharia de atributos é substituído por uma extração automática. Além disso, consideram a dependência temporal entre segmentos adjacentes, demonstrada em [105] com um modelo oculto de Markov (HMM). Por fim, em [111] foi usada uma rede LSTM (memória de longo prazo) que analisa apenas a variabilidade da frequência cardíaca instantânea. Os autores referem que a técnica poderia ser usada nos dispositivos de pulso atuais, que fornecem esta medida.

Por último, como mencionado em [107], seria de interesse avaliar a aplicabilidade de sensores PPG nesta tarefa. Estes possibilitam a análise da FC, respiração e saturação de oxigênio. Contudo, estas medidas são raramente obtidas nos dispositivos de pulso, consideradas imprecisas em comparação à oximetria medida no dedo [112, 113]. Em particular, a última não é possível na maioria das soluções comerciais, pois requer luz específica do espectro.

Ref. Ano	Atributos	No.	Algoritmo	Acc(%)	Se(%)	Sp(%)
[104] 2003	VFC e EDR	88	LDA	90	86.4	92.3
[114] 2015	VFC e EDR	2	Baseado em regras	93.2	95.7	91
[105] 2015	VFC e EDR	6	SVM + HMM	86.2	82.6	88.4
[115] 2016	VFC e complexo QRS por polinômios de Hermite	18	SVM	83.8	79.5	88.4
[116] 2016	VFC e EDR	20	ANN	82.1	88.4	72.3
[111] 2017	FC	60	LSTM-RNN	100	-	-
[107] 2018	VFC e amplitude RS no domínio do tempo	2	SVM	85.7	81.4	88.4
[110] 2018	Obtidos por <i>sparse autoencoder</i>	-	DNN + HMM (<i>decision fusion</i>)	84.7	88.9	82.1
[109] 2019	VFC (3 min)	19	ANN	84	84.5	83.7
[106] 2019	Intervalo RR e amplitude (5 min)	-	CNN	87.6	83.1	90.3
[117] 2019	Intervalo RR	-	LSTM-CNN	99.8	99	96.9

Tabela 2.5: Estudos de detecção de apneia do sono com análise por-segmento. Acc–exatidão; Se–sensibilidade; Sp–especificidade; DNN–rede neuronal profunda.

Ref.	Ano	Algoritmo	Acc(%)	Se(%)	Sp(%)	Corr.
[105]	2015	SVM + HMM	97.1	95.8	100	0.860
[115]	2016	SVM	97.1	100	100	0.841
[110]	2018	DNN + HMM	100	100	100	-
[106]	2019	CNN	97.1	100	91.7	0.943

Tabela 2.6: Estudos de detecção de apneia do sono com análise por-gravação.

Apesar de amplamente utilizado, algoritmos treinados com sucesso neste *dataset* apresentam um desempenho reduzido em outros *datasets*, que incluem pacientes com um espectro mais amplo de eventos apneicos e distúrbios do sono [118]. Portanto, este não abrange a complexidade e variedade do sono necessárias para uma solução que possa ser empregue em situações mais complexas. Outra limitação é a divisão dos dados em segmentos de 1 min, sendo que um evento pode ocorrer no meio de dois segmentos, e um segmento pode conter mais do que um evento [106]. Além disso, não separa eventos de hipopneia e apneia, e os eventos centrais não são incluídos. Assim, torna-se necessário considerar *datasets* adicionais para solucionar estes problemas.

2.8.1 Soluções comerciais

Alguns dispositivos de pulso podem ser usados na detecção de apneia do sono. O relógio Apple Watch tem sensores PPG e mais recentemente ECG, que permite o desenvolvimento de aplicações de terceiros. De forma semelhante, os dispositivos Fitbit contêm sensores de FC e saturação de oxigênio, e disponibilizam uma aplicação que, por enquanto, apenas reporta os valores lidos. Por fim, o ScanWatch [119] foi introduzido como o primeiro relógio com sensor ECG, clinicamente testado. Contém três elétrodos e sensor de saturação de oxigênio, e permite também a detecção de fibrilação atrial e classificação do ritmo cardíaco.

Apesar de várias soluções estarem em desenvolvimento, é de notar que ainda nenhum destes dispositivos obteve certificação, a qual é essencial para um diagnóstico válido.

2.9 Sensores e Datasets

2.9.1 Datasets

- “CAP Sleep Database” [120]: contém 108 exames de polissonografia. Para cada indivíduo, é identificada a existência ou não de patologia, e são disponibilizadas as leituras dos sinais (.edf) e classificação das fases do sono (.txt). Além disso, é identificada a fase A do padrão cíclico alternante (CAP). Este ciclo é dividido nas fases A (A1, A2 e A3) e B consoante a atividade do EEG no sono NREM, e está associado à instabilidade do sono. São incluídos sinais de EEG, EOG, EMG, ECG e de respiração, e as patologias anotadas são as seguintes: bruxismo, insónia, narcolepsia, epilepsia noturna do lobo frontal, movimentos periódicos das pernas, distúrbio comportamental do sono REM e distúrbios respiratórios do sono.
- “Sleep-EDF Database” [121]: Inclui 153 gravações de polissonografia do estudo “Sleep Cassette”. Contém sinais de EEG, EOG, EMG, respiração e temperatura (.edf) e são disponibilizados os hipnogramas correspondentes (.edf+) e informação dos participantes (.xls).
- “EEG/EOG/EMG data from a cross sectional study on psychophysiological insomnia and normal sleep subjects” [122]: contém dados da polissonografia de 22 indivíduos com ou sem diagnóstico de insónia. Inclui os sinais de EEG, EOG, EMG e ECG (.edf) e a classificação das fases do sono (.txt). Além disso, são disponibilizados atributos no domínio da frequência e não linear, e informação demográfica dos participantes (.xlsx).
- “Newcastle polysomnography and accelerometer data” [123]: contém dados de um acelerómetro (.bin) de 28 indivíduos. A classificação do sono é obtida com a polissonografia (.txt) e a informação dos participantes (.csv) inclui o diagnóstico de distúrbios do sono.
- “Motion and heart rate from a wrist-worn wearable and labeled sleep from polysomnography” [124]: contém dados de aceleração e frequência cardíaca de 31 indivíduos, obtidas a partir do sensor PPG de um dispositivo Apple Watch. As fases do sono são classificadas de acordo com a polissonografia (.txt).

2.9.2 Sensores

- Empatica E4 [125]: O E4 é um dispositivo de pulso que permite a análise de dados fisiológicos em tempo real. Inclui sensores de temperatura, AED, PPG e movimento. O acesso aos dados de sessões concluídas é feito através de uma plataforma web. Além disso, é disponibilizada uma aplicação e APIs (do inglês *Application Programming Interface*) para o desenvolvimento de aplicações móveis em Android e iOS. Nesse caso, a execução do E4 em modo de transmissão permite monitorizar os dados através de uma conexão BLE (do inglês *Bluetooth Low Energy*). Para a análise da VFC é possível aceder ao sinal IBI, o qual é obtido a partir do processamento do sinal PPG com um algoritmo que remove falsos picos causados por ruído.
- BITalino [126]: o BITalino é uma plataforma desenvolvida para análise de sinais fisiológicos. O Freestyle Kit inclui sensores como EMG, ECG, AED e EEG, e permite que sejam usados em *wearables*, com impressão de modelos 3D. Através de comunicação por Bluetooth ou BLE, é possível visualizar os dados com o software OpenSignals, assim como processar os mesmos através dos módulos disponibilizados. A exportação dos dados pode ser realizada em vários formatos, compatíveis com plataformas como o Matlab e Python. Além disso, as APIs permitem incorporar o BITalino em aplicações de software.
- MySignals [127]: a plataforma MySignals HW permite desenvolver aplicações para monitorização biométrica. Disponibiliza 17 sensores para medir mais de 20 parâmetros, porém, sem possibilidade de serem usados em simultâneo. Para o desenvolvimento é necessário um Arduino Uno, e através do SDK (do inglês *Software Development Kit*) é possível aceder ao valor dos sensores e transmitir informação por BLE. Ao utilizar a aplicação móvel, os dados podem ser enviados por Wi-Fi para um site, onde são armazenados e acedidos por uma API. No entanto, este serviço implica um plano pago anualmente.
- Everion [128]: O Everion é um equipamento médico validado que combina sensores precisos numa única plataforma. É utilizado no braço e mede 22 parâmetros de forma contínua e não invasiva. A solução de monitorização compreende o armazenamento dos dados, funcionalidades analíticas e visualização. Porém, também permite integração com API e SDK.
- Polar [129]: Os sensores Polar são amplamente reconhecidos pela sua precisão na monitorização da frequência cardíaca. O Polar H10 mede o ECG através de um banda colocada no peito, enquanto que o Polar OH1 têm sensores PPG e é utilizado no braço. Além disso, disponibilizam dados de aceleração. O SDK permite o desenvolvimento de aplicações móveis para ler e interpretar os dados em tempo real, transmitidos por Bluetooth.
- PulseOn OHR Tracker [130]: o dispositivo OHR Tracker é utilizado no pulso e permite uma fácil integração através de API, SDK e aplicação Android. Os dados são transmitidos por Bluetooth, e é possível aceder aos valores de aceleração e do sensor PPG, assim como dados processados de FC e IBI, para a análise da VFC.

2.9.3 Discussão

O estudo em [131] avalia a qualidade da FC e métricas da VFC obtidas de dispositivos usáveis com tecnologia PPG, comparativamente a um ECG. Foi concluído que, apesar de adequados para a monitorização, estes podem ser limitados em condições não estacionárias, nomeadamente, durante a atividade física. Além disso, os participantes demonstraram preferência por dispositivos utilizados no pulso do que no braço. No entanto, a posição do pulso pode resultar em medições de baixa qualidade e estes dispositivos são mais suscetíveis a artefactos de movimento. Assim, existe um compromisso entre o conforto e a fiabilidade. Por fim, na análise da FC, o dispositivo Everion obteve melhores resultados em comparação ao E4, e permitiu a análise da VFC.

De forma semelhante, em [132] foi possível concluir que os dispositivos de pulso são precisos e fiáveis para a deteção da FC, quando a quantidade de movimentos da mão não é excessiva. No entanto, a VFC apenas pode ser estimado de forma fiável, a partir de dados IBI, durante condições de repouso. Além disso, ao comparar o E4 e o PulseOn, o segundo obteve maior precisão.

Concluindo, a análise da frequência cardíaca é mais precisa ao medir o ECG. No entanto, é possível obter resultados comparáveis com sensores PPG, principalmente quando os dispositivos são utilizados no braço, e em condições estacionárias.

Capítulo 3

Classificação do sono

3.1 Dados

Neste trabalho foi utilizado o *dataset* “EEG/EOG/EMG data from a cross sectional study on psychophysiological insomnia and normal sleep subjects” [122], previamente abordado na subsecção 2.9.1. Este consiste em 22 gravações com duração aproximada de 8h, na qual os indivíduos pertencem à faixa etária 18–63. A frequência de amostragem do ECG é de 256 Hz, com unidades de microvolt. A tabela 3.1 apresenta informação sobre a composição do mesmo. Nas experiências realizadas apenas foram considerados os indivíduos sem insónia.

Para o processamento e tratamento dos dados, foi usada a linguagem de programação Python e as bibliotecas pyhrv [133], scikit-learn [134] e keras [135].

Condição	Vigília	REM	N1	N2	N3+N4	Total
Normal (11)	1272 (12%)	708 (7%)	3749 (36%)	2286 (22%)	2379+26 (23%)	10420
Insónia (11)	3285 (33%)	1140 (11%)	2897 (29%)	1341 (13%)	1443+10 (14%)	10116

Tabela 3.1: Distribuição dos segmentos por fase do sono. As fases N3 e N4 foram concatenadas pois atualmente não existe distinção entre ambas.

3.1.1 Pré-processamento

Inicialmente, o sinal de ECG foi sincronizado no tempo com os resultados da polissonografia, considerando segmentos de 30 s. Para a análise da VFC, são recomendadas janelas temporais de 2 a 5 min [61]. Assim, com o objetivo de analisar o impacto no desempenho, o sinal foi dividido em segmentos de 1.5 min, 2.5 min, 3.5 min, e 4.5 min, centrados em cada intervalo de 30 s. A figura 3.1 ilustra esta abordagem para a menor dimensão. Por último, os segmentos identificados como movimento foram removidos.

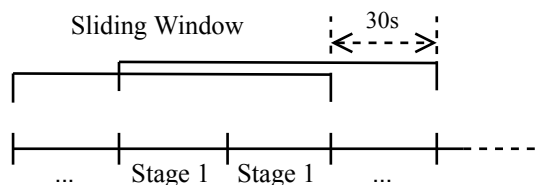


Figura 3.1: Abordagem para processamento dos segmentos.

Para obter a série de intervalos entre batimentos (RR), foi necessário detetar os picos do ECG. O exemplo de um segmento de 10 s é apresentado na figura 3.2. Considerando que se trata de um cálculo sensível a ruído, segmentos com secções de movimento podem pôr em causa o seu rigor. Além disso, a fiabilidade da classificação é indicada no *dataset*, pelo que será analisado o impacto de ambos os aspetos no desempenho.

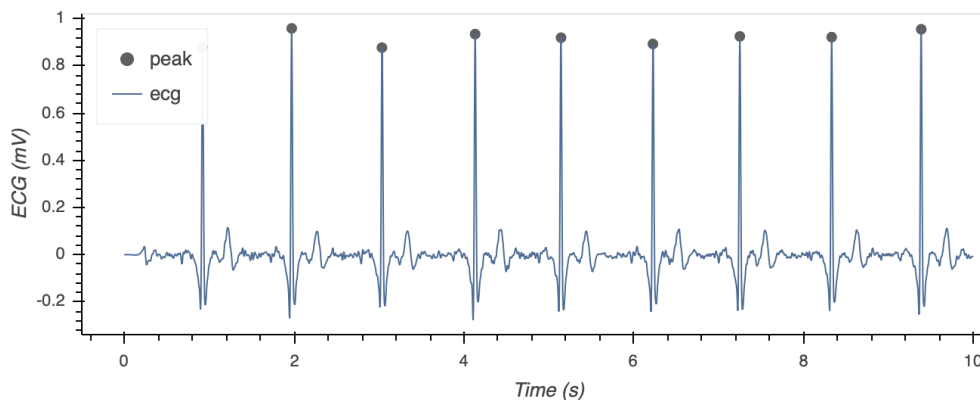


Figura 3.2: Exemplo de um segmento do eletrocardiograma.

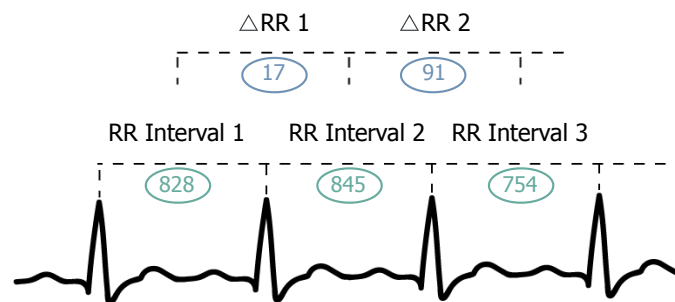
3.1.2 Extração de atributos

Os segmentos foram processados no domínio do tempo, frequência e não linear. Os 34 atributos obtidos são apresentadas na tabela 3.2, e brevemente abordados de seguida.

Para a análise no domínio do tempo, são usadas as séries RR e Δ RR, onde a segunda representa a diferença entre intervalos adjacentes como ilustrado na figura 3.3. No domínio da frequência, para o cálculo do PSD foi aplicado o método Lomb-Scargle. Esta é uma técnica não-paramétrica, que não requer informação prévia dos parâmetros do sinal ou interpolação quando o sinal é não-periódico [61]. Assim, foram consideradas três bandas: VLF (0–0.04 Hz), LF (0.04–0.15 Hz) e HF (0.15–0.4 Hz). Por último, no domínio não linear, foi utilizado o diagrama de Poincaré, que permite visualizar cada intervalo RR em relação ao anterior, o método DFA (do inglês *Detrended Fluctuation Analysis*), usado na análise de correlação em séries temporais, e calculada a entropia, que avalia a sua complexidade.

Domínio	No.	Parâmetros
Tempo	14	
Série RR	5	Número (<i>nni_counter</i>), média (<i>nni_mean</i>), máximo (<i>nni_max</i>), mínimo (<i>nni_min</i>), desvio padrão (<i>sdnn</i>)
Série Δ RR	9	Média (<i>diff_mean</i>), máximo (<i>diff_max</i>), mínimo (<i>diff_min</i>), desvio padrão (<i>sdsd</i>), valor quadrático médio (<i>rmsd</i>), número e percentagem de valores maiores que 50 (<i>nn50</i> , <i>pp50</i>) e maiores que 20 (<i>nn20</i> , <i>pp20</i>)
Frequência	13	
Para cada banda	9	Frequência pico (<i>peak_vlf</i> , <i>peak_lf</i> , <i>peak_hf</i>), potência absoluta (<i>abs_vlf</i> , <i>abs_lf</i> , <i>abs_hf</i>), potência relativa (<i>rel_vlf</i> , <i>rel_lf</i> , <i>rel_hf</i>)
Outros	4	Potência total (<i>lomb_total</i>), LF/HF (<i>lomb_ratio</i>), potência normalizada das bandas LF e HF (<i>norm_lf</i> , <i>norm_hf</i>)
Não linear	7	
Diagrama de Poincaré	4	Desvio padrão a curto- (<i>sd1</i>) e longo-prazo (<i>sd2</i>), área da elipse (<i>ellipse_area</i>), SD2/SD1 (<i>sd_ratio</i>)
Análise adicional	3	Entropia amostral (<i>sampen</i>), DFA a curto- (<i>dfa_alpha1</i>) e longo-prazo (<i>dfa_alpha2</i>)

Tabela 3.2: Atributos calculados do eletrocardiograma.

Figura 3.3: Séries obtidas do eletrocardiograma: intervalo entre batimentos (RR) e diferença entre intervalos consecutivos (Δ RR), expressas em ms.

3.2 Algoritmos

Os métodos de aprendizagem supervisionada são frequentemente utilizados em problemas de classificação. Para tal, os modelos são treinados através do valor de referência, de modo aproximar uma função que mapeia os valores de entrada e saída. Dependendo da aplicação, cada classificador tem pontos fortes e fracos mas, em última análise, o seu desempenho depende dos dados. Como tal, a melhor abordagem é avaliar diferentes modelos.

De seguida são descritos os algoritmos utilizados neste trabalho.

3.2.1 Support Vector Machine

O SVM (do inglês *Support Vector Machine*) é um classificador linear binário. O seu objetivo é encontrar um hiperplano num espaço N -dimensional, sendo N o número de atributos, que classifique distintamente os dados, representados por pontos. Nesta divisão, o plano contém a margem máxima, ou seja, a maior distância entre os pontos de ambas as classes, como ilustrado em 3.4.

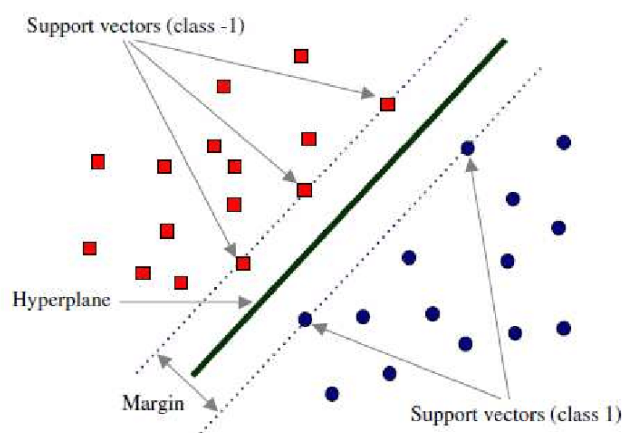


Figura 3.4: Hiperplano de classificação do algoritmo *Support Vector Machine*. Fonte: [136]

Apesar de simples, a maioria dos casos requer considerações adicionais, nomeadamente, se os dados não são linearmente separáveis, e em problemas multi-classe. No primeiro caso, é possível utilizar uma *soft margin*, que tolera um certo número de erros na classificação, ou mapear os dados para um espaço dimensional superior, através de funções *kernel*. No segundo, a conversão para o formato binário é obtida com as abordagens *one-vs-one* ou *one-vs-rest*.

3.2.2 Análise Discriminante Linear

A análise discriminante linear (LDA) é uma técnica de redução de dimensionalidade frequentemente utilizada em problemas de classificação. Para a separação dos dados, estes são projetados num subespaço dimensional inferior, necessariamente menor que o número de classes. Um exemplo deste método é apresentado na figura 3.5. O critério utilizado visa maximizar a distância entre a média das classes, e minimizar a variação de cada classe.

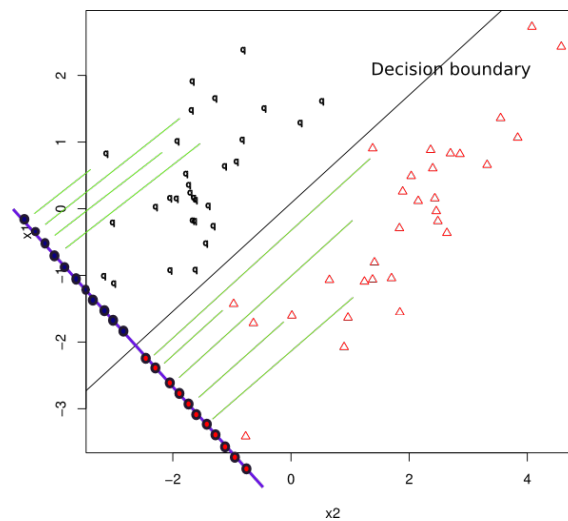


Figura 3.5: Aplicação da análise discriminante linear num problema de 2 classes e 2 dimensões: os pontos são projetados num espaço de 1 dimensão. Fonte: [137]

Para ser utilizado como classificador, o LDA assume que a distribuição condicional das classes é gaussiana, com a mesma matriz de covariância. A análise quadrática (QDA) é uma generalização deste método, sem a segunda suposição.

3.2.3 Random Forest

O RF (do inglês *Random Forest*) é um método *ensemble*, onde são gerados vários classificadores do mesmo tipo e os seus resultados são combinados. As árvores de decisão são consideradas instáveis, ou seja, bastante sensíveis à variação dos dados de treino. Por conseguinte, cada modelo individual considera apenas um número pré-definido de atributos em cada nó, e através de uma abordagem *bagging*, treina com um *subset* aleatório. Ao aumentar a sua diversidade é obtido um melhor desempenho preditivo. A arquitetura descrita é apresentada de seguida, na figura 3.6.

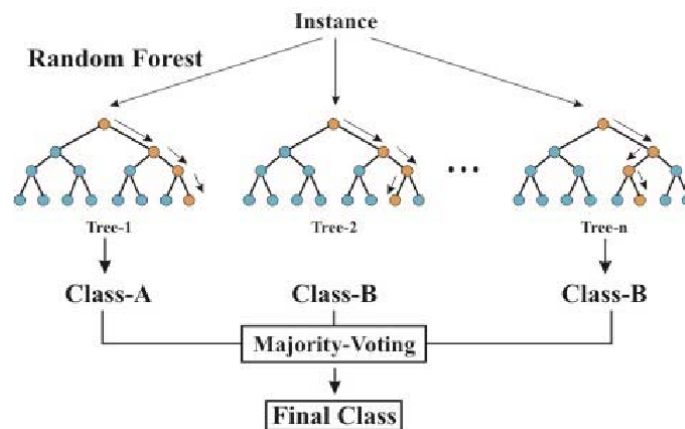


Figura 3.6: Arquitetura do algoritmo *Random Forest*. Fonte: [138]

3.2.4 K-Nearest Neighbors

O KNN (do inglês *K-Nearest Neighbors*) classifica uma instância de acordo com a classe de instâncias similares, sem a criação explícita de um modelo. Para tal, efetua a votação majoritária das k instâncias de treino mais próximas. Os desafios deste classificador incluem a velocidade da previsão, a escolha do hiperparâmetro k e da função de cálculo das distâncias, e o efeito do peso em função das dimensões. A ilustração do processo de classificação deste algoritmo é mostrada a seguir, na figura 3.7.

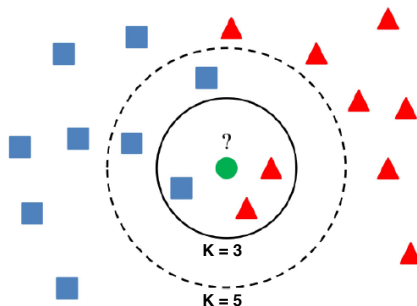


Figura 3.7: Aplicação do algoritmo *k-Nearest Neighbors*. O círculo verde é classificado consoante os k vizinhos mais próximos. Adaptado de: [139]

3.2.5 Rede Neuronal Recorrente

As redes neurais artificiais são fundamentadas em estudos sobre a estrutura do cérebro humano, na tentativa de reproduzir a sua forma inteligente de processar informação. Estas são compostas por várias unidades de processamento organizados em camadas, e um conjunto de elos de ligação com pesos associados. Para a aprendizagem, é geralmente usado o algoritmo *backpropagation*, de forma a atualizar os pesos da rede consoante o erro obtido.

As redes *feedforward* têm uma estrutura onde os dados circulam numa única direção, partindo da camada de entrada, para as camadas intermédias e camada de saída. Contudo, como não possuem memória dos *inputs* recebidos previamente, não são ideais em situações que envolvam dados sequenciais. Assim, as RNN abordam este problema ao permitir que a informação persista, como é ilustrado na figura 3.8. Estas têm algumas variantes como as redes LSTM, criadas explicitamente para aprender dependências de longo prazo.

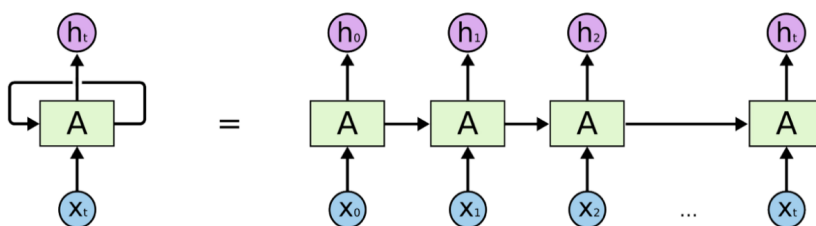


Figura 3.8: Arquitetura básica de uma rede neuronal recorrente. A imagem à esquerda indica a sua forma simplificada. A variável x_t representa o input no tempo t , e h_t o output. Fonte: [140]

3.3 Validação

3.3.1 Validação cruzada

A validação em *machine learning* é o processo de avaliar um modelo treinado num conjunto de dados de teste de modo a estimar a sua capacidade de generalização em novos contextos. Assim, permite identificar situações como o sobreajuste (do inglês *overfitting*), na qual um modelo está demasiado adaptado aos dados de treino. Geralmente os dados são divididos em três partes: treino, validação e teste. Pela mesma ordem, estes são usados para o treino do modelo, ajuste dos hiperparâmetros, e na avaliação imparcial do modelo final. Este método de separação é denominado de validação cruzada *holdout*.

A validação cruzada *k-fold* é uma estratégia que consiste na divisão do conjunto de dados em k subconjuntos, onde $k-1$ divisões são usadas para o treino do modelo e a divisão sobranete para teste. O processo é repetido iterativamente, como ilustrado na figura 3.9. No fim, calcula-se a média das métricas de desempenho. Este método pode ser considerado em casos específicos, nomeadamente *leave-one-out*, com k igual ao número total de dados, e *leave-one-subject-out*, de modo a garantir que os subconjuntos são compostos por diferentes indivíduos. Por fim, o processo de estratificação considera a distribuição original das classes, diminuindo a variância da estimativa.

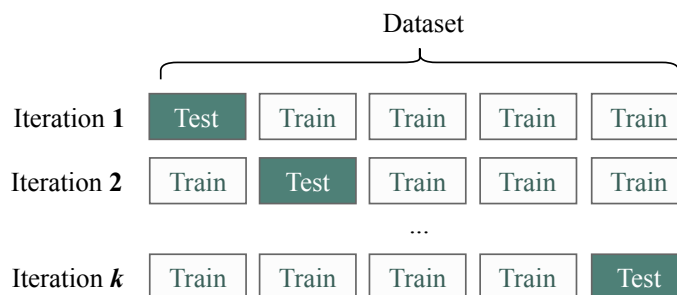


Figura 3.9: Validação cruzada k -fold.

Como a questão da variabilidade individual é crucial nos sistemas de classificação do sono, foi considerada validação *leave-one-subject-out*. Assim, em cada iteração, são usados 10 indivíduos para treino e um para teste. Adicionalmente, foi aplicada validação estratificada *10-fold*. Tendo em conta o desequilíbrio de classes, é garantida a mesma proporção de exemplos em cada subconjunto.

3.3.2 Dados não balanceados

Um conjunto de dados é considerado não balanceado quando difere significativamente no número de instâncias entre classes. Este desequilíbrio tem como consequência modelos enviesados, que beneficiam a classe maioritária, produzindo por isso estimativas pouco fiáveis.

Atualmente existem vários métodos capazes de contornar este problema. As técnicas de amostragem consideram a remoção ou adição de instâncias de modo a obter uma distribuição balanceada, nomeadamente ao aplicar sub-amostragem (do inglês *under-sampling*) e sobre-amostragem

(do inglês *over-sampling*), como demonstrado na figura 3.10. Na primeira, as observações são selecionadas aleatoriamente ou é aplicado um algoritmo de *clustering*, com o objetivo de preservar a informação subjacente à classe majoritária. Por outro lado, na segunda são criadas observações da classe minoritária. [141]

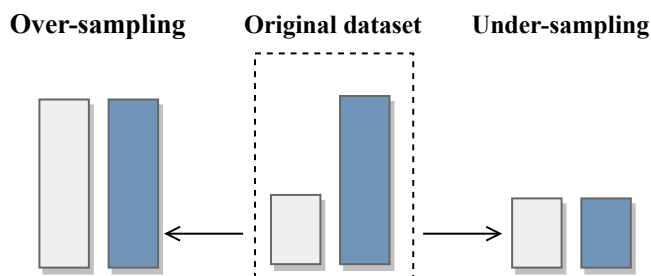


Figura 3.10: Aplicação de sub-amostragem e sobre-amostragem.

As técnicas de sobre-amostragem incluem o ROS (do inglês *Random Over-Sampling*), e o SMOTE (do inglês *Synthetic Minority Over-Sampling Technique*). O ROS consiste na cópia exata de instâncias existentes, selecionadas de forma aleatória. No sentido de diminuir a probabilidade de sobreajuste introduzida por este método, o SMOTE insere instâncias sintéticas. Para tal, procura os k vizinhos mais próximos para cada instância da classe minoritária, e seleciona um de forma aleatória. De seguida, obtém a diferença entre os dois, a qual é multiplicada por um valor aleatório compreendido entre 0 e 1, e adicionada ao inicial. Um exemplo deste funcionamento é apresentado de seguida na figura 3.11.

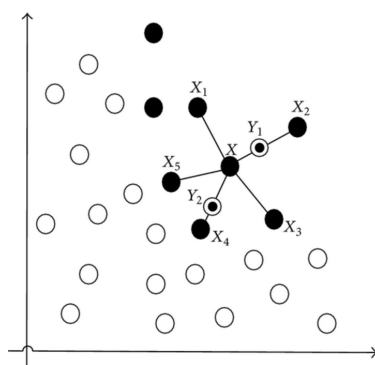


Figura 3.11: Aplicação do algoritmo SMOTE, com $k = 5$. Os vizinhos mais próximos de X são X_1 a X_5 , e Y_1 e Y_2 as novas instâncias. Fonte: [142]

Em termos do *dataset* utilizado, a classe REM é quase inexistente. Desta forma, a abordagem de sub-amostragem não foi considerada, pois as restantes classes seriam diminuídas severamente. Portanto, para ajustar o número de instâncias entre classes, foi usada uma variação do SMOTE denominada de SVM-SMOTE. O algoritmo de classificação permite uma melhor delimitação entre classes, onde as instâncias próximas das fronteiras são mais relevantes [143].

3.3.3 Seleção de atributos

A seleção do conjunto adequado de atributos a ser utilizado em modelos de *machine learning* permite melhorar o seu desempenho, e reduzir o custo computacional. De acordo com a estratégia de pesquisa, os algoritmos de *feature selection* podem ser divididos em três categorias: *filter*, *wrapper* e *embedded*. Se o critério de seleção é independente do classificador, caracterizam-se como sendo do tipo *filter*, e são baseados na informação estatística dos dados. Por outro lado, os métodos *wrapper* consideram o desempenho do classificador quando sujeito a diferentes subconjuntos, selecionados de forma iterativa. Por fim, os métodos *embedded* consistem numa combinação dos anteriores, ao incorporar a seleção na construção do modelo. [144]

Os métodos *filter* tendem a ser rápidos e robustos ao sobreajuste. Assim, com o objetivo de eliminar atributos irrelevantes, as dependências entre variáveis foram avaliadas através do coeficiente de correlação de Pearson.

3.3.4 Medidas de desempenho

Ao treinar um determinado modelo, é importante avaliar o seu desempenho. Neste sentido, um dos critérios mais comuns é a exatidão (*ACC*), que indica a proporção de exemplos corretamente classificados (equação 3.1). Porém, este não é ideal em problemas de dados não balanceados, uma vez que não distingue o número correto de classificações das instâncias de classes diferentes.

$$ACC = \frac{VP + VN}{VP + VN + FP + FN} \quad (3.1)$$

A matriz de confusão possibilita a visualização do desempenho do modelo, como representado na tabela 3.7. Esta consiste numa matriz quadrada de ordem igual ao número de classes, onde cada linha contém as instâncias com a sua classe original, e cada coluna com a classe prevista.

		Previsão	
		Positivo	Negativo
Verdade	Positivo	Verdadeiro positivo (<i>VP</i>)	Falso negativo (<i>FN</i>)
	Negativo	Falso positivo (<i>FP</i>)	Verdadeiro negativo (<i>VN</i>)

Tabela 3.3: Exemplo de uma matriz de confusão para um problema binário.

A taxa de verdadeiros positivos (equação 3.2) conhecida como sensibilidade (*SENS*) ou *recall*, indica a proporção de instâncias positivas corretamente identificadas.

$$SENS = \frac{VP}{VP + FN} \quad (3.2)$$

A taxa de verdadeiros negativos (equação 3.3), referida como especificidade (SPEC), indica a proporção de instâncias negativas corretamente identificadas.

$$SPEC = \frac{VN}{VN + FP} \quad (3.3)$$

O valor preditivo positivo (equação 3.4) designado de precisão (PREC), indica a proporção de instâncias positivas corretamente identificadas no conjunto de instâncias previstas como positivas.

$$PREC = \frac{VP}{VP + FP} \quad (3.4)$$

O F_1 Score (equação 3.5) é uma medida que considera a sensibilidade e precisão.

$$F_1 = 2 \times \frac{PREC \times SENS}{PREC + SENS} \quad (3.5)$$

O coeficiente Kappa de Cohen (equação 3.6) é uma medida da concordância entre avaliadores, onde p_o representa a proporção observada de concordâncias, e p_e a proporção esperada.

$$k = \frac{p_o - p_e}{1 - p_e} \quad (3.6)$$

Por fim, a curva característica de operação do receptor (ROC) é um critério de avaliação gráfico do desempenho da classificação, que permite o estudo da variação da sensibilidade e especificidade para vários valores de corte. A área sob a curva (AUC) é calculada de forma a comparar diferentes modelos. Quanto maior for este valor, melhor é o desempenho, como ilustrado na figura 3.12.

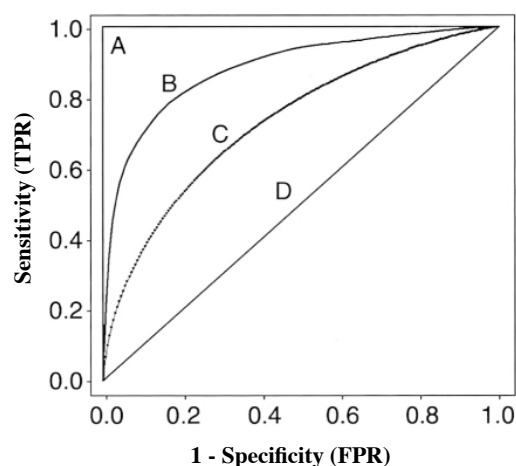


Figura 3.12: Comparação de curvas ROC com diferentes valores de área sob a curva. Os testes com capacidade de distinção estão entre os extremos A e D. Adaptado de: [145]

3.4 Resultados

A abordagem considerada aplica dois tipos de validação através de um teste independente (UIT) e dependente do indivíduo (UDT). Nomeadamente, validação cruzada estratificada 10-*fold*, e *leave-one-subject-out*, doravante referenciadas por 10-VC e LOSO-VC.

Em cada iteração, os dados são normalizados a nível dos atributos, com uma média de zero e desvio padrão de um. De seguida, após a seleção do melhor subconjunto de atributos, é efetuada a sobre-amostragem do conjunto de treino. Esta ordem é justificada pela dependência entre variáveis introduzida pela última [146]. Por fim, as medidas de desempenho são calculadas como a média de todas as iterações, em particular a exatidão e a sensibilidade das classes. A figura 3.13 ilustra o processo descrito.

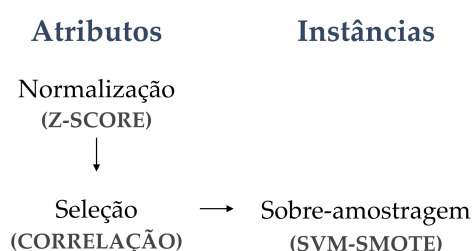


Figura 3.13: Ilustração da abordagem utilizada.

Foram avaliados os algoritmos SVM, KNN, LDA e RF com os seus hiper-parâmetros ajustados manualmente. Além do desempenho, é analisado o impacto dos seguintes aspetos:

- dimensão dos segmentos (de 1.5 min a 4.5 min);
- classificação de três ou quatro classes;
- filtragem de segmentos com movimento, ou identificados com fiabilidade menor que 35%.

3.4.1 Resultados da classificação

Inicialmente foi avaliada a classificação de 4 classes, ao variar a dimensão dos segmentos. Os resultados do teste independente do indivíduo são apresentados de seguida, nas tabelas 3.4 e 3.5. Da sua análise, é possível observar que com uma janela de 1.5 min não existe discriminação clara entre classes. Com o seu aumento, os algoritmos KNN e RF apresentam uma melhoria substancial, nomeadamente o RF, que atinge uma exatidão de 78%. Assim, a maior dimensão dos segmentos permite obter um maior contexto, sem sobrecarga de informação. Por fim, é de notar que a classe maioritária N1+2 apresenta a menor sensibilidade, e é frequentemente classificada como N3.

Ao mudar para um teste dependente do indivíduo, o desempenho diminuiu bastante. Nomeadamente, o RF com segmentos de 4.5 min, apenas atinge uma exatidão de 51% com sensibilidade à fase N3 de 19%, N1+2 de 75%, REM de 20% e vigília de 28%. Além do mais, os resultados têm uma grande variação. Portanto, conclui-se que a abordagem usada não permite classificar o sono com precisão neste contexto.

UIT (10-VC)				
Modelo	1.5 min	2.5 min	3.5 min	4.5 min
SVM	0.55 (0.01)	0.57 (0.01)	0.58 (0.02)	0.59 (0.02)
LDA	0.46 (0.02)	0.46 (0.01)	0.47 (0.01)	0.46 (0.01)
KNN	0.46 (0.01)	0.49 (0.02)	0.52 (0.01)	0.55 (0.03)
RF	0.65 (0.01)	0.72 (0.01)	0.76 (0.01)	0.78 (0.01)

Tabela 3.4: Resultados da classificação de quatro classes (exatidão e desvio padrão), variando os algoritmos e dimensão dos segmentos.

UIT (10-VC)				
Modelo	1.5 min	2.5 min	3.5 min	4.5 min
SVM	D: 0.46 (0.05)	D: 0.55 (0.06)	D: 0.62 (0.03)	D: 0.62 (0.04)
	L: 0.57 (0.03)	L: 0.56 (0.03)	L: 0.54 (0.02)	L: 0.53 (0.04)
	R: 0.4 (0.04)	R: 0.49 (0.05)	R: 0.57 (0.07)	R: 0.65 (0.06)
	W: 0.69 (0.06)	W: 0.71 (0.05)	W: 0.74 (0.04)	W: 0.76 (0.04)
LDA	D: 0.54 (0.05)	D: 0.57 (0.03)	D: 0.59 (0.04)	D: 0.61 (0.03)
	L: 0.4 (0.03)	L: 0.38 (0.02)	L: 0.39 (0.02)	L: 0.35 (0.02)
	R: 0.3 (0.04)	R: 0.34 (0.06)	R: 0.42 (0.08)	R: 0.5 (0.06)
	W: 0.68 (0.05)	W: 0.68 (0.05)	W: 0.66 (0.04)	W: 0.64 (0.05)
KNN	D: 0.53 (0.03)	D: 0.6 (0.06)	D: 0.65 (0.03)	D: 0.64 (0.04)
	L: 0.39 (0.03)	L: 0.4 (0.04)	L: 0.41 (0.02)	L: 0.43 (0.04)
	R: 0.42 (0.07)	R: 0.56 (0.08)	R: 0.66 (0.07)	R: 0.74 (0.05)
	W: 0.67 (0.06)	W: 0.69 (0.06)	W: 0.75 (0.05)	W: 0.79 (0.05)
RF	D: 0.49 (0.03)	D: 0.64 (0.03)	D: 0.68 (0.03)	D: 0.71 (0.03)
	L: 0.73 (0.02)	L: 0.75 (0.02)	L: 0.78 (0.02)	L: 0.81 (0.02)
	R: 0.42 (0.05)	R: 0.59 (0.05)	R: 0.69 (0.05)	R: 0.75 (0.04)
	W: 0.7 (0.06)	W: 0.77 (0.04)	W: 0.81 (0.04)	W: 0.83 (0.04)

Tabela 3.5: Resultados detalhados da classificação de quatro classes (sensibilidade e desvio padrão de cada fase do sono), variando os algoritmos e dimensão dos segmentos.

De seguida foi considerada a classificação de 3 classes, com os resultados descritos nas tabelas 3.6 e 3.8. Comparativamente aos resultados anteriores, observa-se um aumento geral da exatidão. O tamanho dos segmentos continua sem impacto positivo no algoritmo LDA, no entanto, permite uma maior distinção da fase REM. O RF obteve novamente o melhor desempenho, que é possível analisar em detalhe na matriz de confusão descrita na tabela 3.7.

Em relação ao teste dependente do indivíduo, a melhoria dos resultados não é ainda suficiente para combater a variabilidade individual. Existe uma baixa sensibilidade às fases do sono, como acentuado pelo algoritmo RF na fase REM. Este problema pode também ser justificado pela grande diferença da distribuição de classes entre indivíduos. Em particular, a fase REM varia de 1 a 118 instâncias, e a vigília de 7 a 331. Assim, para melhorar esta componente, seria necessária a análise de um conjunto de dados mais abrangente.

UIT (10-VC)				
Modelo	1.5 min	2.5 min	3.5 min	4.5 min
SVM	0.75 (0.01)	0.77 (0.01)	0.79 (0.01)	0.79 (0.01)
LDA	0.74 (0.01)	0.72 (0.01)	0.71 (0.01)	0.7 (0.01)
KNN	0.69 (0.02)	0.7 (0.02)	0.72 (0.01)	0.73 (0.01)
RF	0.84 (0.01)	0.87 (0.0)	0.89 (0.01)	0.9 (0.01)
UDT (LOSO-VC)				
Modelo	1.5 min	2.5 min	3.5 min	4.5 min
SVM	0.69 (0.1)	0.67 (0.17)	0.68 (0.18)	0.67 (0.19)
LDA	0.63 (0.07)	0.58 (0.26)	0.6 (0.21)	0.57 (0.24)
KNN	0.63 (0.09)	0.56 (0.14)	0.56 (0.13)	0.55 (0.14)
RF	0.74 (0.12)	0.75 (0.12)	0.75 (0.14)	0.76 (0.14)

Tabela 3.6: Resultados da classificação de três classes (exatidão e desvio padrão), variando os algoritmos e dimensão dos segmentos.

		Previsão		
		NREM	REM	Wake
Verdade	NREM	0.93 (0.01)	0.02 (0.01)	0.05 (0.01)
	REM	0.24 (0.06)	0.71 (0.07)	0.05 (0.02)
	Wake	0.18 (0.04)	0.01 (0.01)	0.81 (0.05)

Tabela 3.7: Matriz de confusão obtida no UIT (10-VC) com o algoritmo *Random Forest* (RF) e segmentos de 4.5 min.

Por fim, relativamente à seleção de segmentos, nas experiências descritas as ambas as condições não foram consideradas. No entanto, a sua aplicação não provocou um impacto significativo no desempenho. O número e percentagem de segmentos removidos é apresentado na tabela 3.9. A restrição da fiabilidade eliminou principalmente instâncias da classe NREM, que constitui a classe maioritária. Da mesma forma, a restrição do movimento não demonstra introduzir ruído na classificação. Assim, torna-se uma análise dispensável, criando oportunidade para uma implementação mais robusta em condições reais de utilização.

Em conclusão, o algoritmo RF obteve o melhor desempenho na distinção das fases do sono comparativamente aos restantes algoritmos. Para esse efeito, a dimensão dos segmentos constituiu um fator relevante, sendo que a sua desvantagem está no maior custo computacional que implica. Contudo, a influência da variabilidade individual foi perceptível, o que não permitiu considerar um cenário dependente do indivíduo.

UIT (10-VC)				
Modelo	1.5 min	2.5 min	3.5 min	4.5 min
SVM	N: 0.79 (0.02)	N: 0.8 (0.02)	N: 0.81 (0.02)	N: 0.8 (0.01)
	R: 0.43 (0.07)	R: 0.54 (0.06)	R: 0.61 (0.04)	R: 0.65 (0.08)
	W: 0.69 (0.04)	W: 0.72 (0.04)	W: 0.76 (0.03)	W: 0.79 (0.04)
LDA	N: 0.78 (0.01)	N: 0.75 (0.02)	N: 0.74 (0.02)	N: 0.72 (0.01)
	R: 0.31 (0.06)	R: 0.42 (0.07)	R: 0.45 (0.05)	R: 0.52 (0.07)
	W: 0.69 (0.05)	W: 0.7 (0.05)	W: 0.69 (0.02)	W: 0.67 (0.02)
KNN	N: 0.7 (0.03)	N: 0.71 (0.02)	N: 0.71 (0.02)	N: 0.71 (0.01)
	R: 0.47 (0.05)	R: 0.61 (0.08)	R: 0.7 (0.04)	R: 0.74 (0.07)
	W: 0.7 (0.06)	W: 0.73 (0.04)	W: 0.78 (0.04)	W: 0.82 (0.03)
RF	N: 0.9 (0.01)	N: 0.91 (0.01)	N: 0.92 (0.01)	N: 0.93 (0.01)
	R: 0.41 (0.05)	R: 0.58 (0.06)	R: 0.68 (0.07)	R: 0.71 (0.07)
	W: 0.7 (0.04)	W: 0.77 (0.04)	W: 0.8 (0.03)	W: 0.81 (0.05)
UDT (LOSO-VC)				
Modelo	1.5 min	2.5 min	3.5 min	4.5 min
SVM	N: 0.81 (0.06)	N: 0.76 (0.19)	N: 0.77 (0.19)	N: 0.76 (0.21)
	R: 0.15 (0.1)	R: 0.26 (0.28)	R: 0.26 (0.27)	R: 0.28 (0.26)
	W: 0.35 (0.23)	W: 0.32 (0.25)	W: 0.31 (0.24)	W: 0.32 (0.26)
LDA	N: 0.72 (0.06)	N: 0.62 (0.29)	N: 0.63 (0.23)	N: 0.6 (0.25)
	R: 0.13 (0.13)	R: 0.46 (0.36)	R: 0.54 (0.33)	R: 0.53 (0.36)
	W: 0.51 (0.3)	W: 0.35 (0.28)	W: 0.37 (0.29)	W: 0.4 (0.31)
KNN	N: 0.73 (0.05)	N: 0.62 (0.17)	N: 0.61 (0.16)	N: 0.6 (0.17)
	R: 0.22 (0.16)	R: 0.32 (0.24)	R: 0.32 (0.26)	R: 0.32 (0.27)
	W: 0.4 (0.18)	W: 0.35 (0.17)	W: 0.43 (0.25)	W: 0.37 (0.19)
RF	N: 0.89 (0.05)	N: 0.9 (0.05)	N: 0.88 (0.11)	N: 0.9 (0.1)
	R: 0.05 (0.05)	R: 0.07 (0.07)	R: 0.15 (0.18)	R: 0.18 (0.23)
	W: 0.33 (0.21)	W: 0.33 (0.18)	W: 0.27 (0.23)	W: 0.25 (0.22)

Tabela 3.8: Resultados detalhados da classificação de três classes (sensibilidade e desvio padrão de cada fase do sono), variando os algoritmos e dimensão dos segmentos.

MOV	RLB	1.5 min	2.5 min	3.5 min	4.5 min
1	1	2501 (23%)	2695 (25%)	2874 (27%)	3061 (29%)
1	0	549 (5%)	799 (8%)	1031 (10%)	1256 (12%)
0	1	2311 (22%)	2300 (22%)	2286 (22%)	2281 (22%)
0	0	253 (2%)	245 (2%)	240 (2%)	239 (2%)

Tabela 3.9: Número e porcentagem de segmentos removidos com diferentes condições. A restrição de movimento e fiabilidade são representadas por MOV e RLB, onde 1 indica o seu uso.

3.4.2 Atributos selecionados

De modo a remover atributos correlacionados, foi definido um *threshold* de 90%. Esta abordagem é considerada inicial pois não avalia correlações com a variável objetivo, nem está otimizada para algum algoritmo. Assim, foi apenas usada com a finalidade de remover atributos irrelevantes, e não afetar a classificação. Foram selecionados cerca de 10 variáveis, sendo que os sub-conjuntos apenas variam ligeiramente ao alterar a dimensão dos segmentos ou o tipo de validação. Entre eles, é observado que *sd1*, *sd2*, *sdsd* e *rmssd* estão correlacionadas com *sdnn*.

Relativamente aos atributos selecionados pelo RF, é possível verificar que o domínio do tempo é considerado o mais relevante. Em contrapartida, os atributos no domínio não linear são ordenados com a menor importância. Assim, com o objetivo de uma solução com menor custo computacional, foram selecionados 7 atributos (*nni_counter*, *nni_mean*, *nni_min*, *nni_max*, *diff_mean*, *diff_max*, *sdnn*). Deste modo, o tempo de execução foi reduzido de cerca de 0.15 s para 0.004 s por segmento, ou seja, em 97%. O impacto no desempenho é analisado de seguida com os resultados na tabela 3.10, mantendo a configuração da classificação de 3 classes com segmentos de 4.5 min, e validação independente do indivíduo.

UIT (10-VC)		
Modelo	Exatidão	Sensibilidade
SVM	0.69 (0.02)	N: 0.67 (0.02) R: 0.7 (0.05) W: 0.8 (0.03)
LDA	0.53 (0.01)	N: 0.55 (0.02) R: 0.49 (0.06) W: 0.43 (0.07)
KNN	0.8 (0.01)	N: 0.79 (0.02) R: 0.79 (0.07) W: 0.86 (0.03)
RF	0.9 (0.01)	N: 0.93 (0.01) R: 0.76 (0.07) W: 0.8 (0.03)

Tabela 3.10: Resultados da classificação de três classes com um conjunto reduzido de atributos no domínio do tempo, e segmentos de 4.5 min.

Ao comparar com os resultados obtidos anteriormente, observa-se que a exatidão do algoritmo SVM e LDA diminui. No entanto, no RF mantém-se, e no KNN aumenta em sete pontos percentuais. Deste modo, é possível concluir que, como o KNN é sensível a *outliers* e atributos irrelevantes, o aumento no desempenho indica que estes possivelmente existem nos restantes atributos. O RF é mais robusto a estes aspetos, e apresenta assim a melhor distinção. Portanto, é considerado uma opção adequada para este problema.

3.4.3 Análise sequencial

Os algoritmos de classificação previamente aplicados não têm em conta o contexto temporal. Dado que a arquitetura do sono apresenta padrões sequenciais, estas abordagens podem não atingir o desempenho ideal, pois não exploram a dependência entre etapas de tempo. Assim, estudos mais recentes investigam modelos que considerem um contexto temporal amplo, principalmente devido à variação do sono à medida que a noite progride.

De modo a avaliar o impacto no desempenho, foi utilizada uma RNN com arquitetura BLSTM, ilustrada na figura 3.14. Neste modelo, cada LSTM têm 8 unidades, o que constitui um total de 16 saídas por cada *timestep*. A função de ativação *softmax* foi usada para a distinção entre classes, com a função de perda *categorical cross-entropy*.

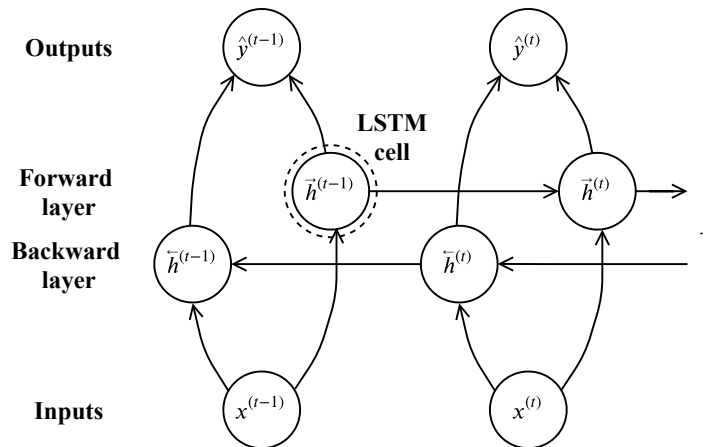


Figura 3.14: Ilustração da arquitetura de redes com memória bidirecional de longo prazo. Adaptado de: [84]

Estas redes são consideradas uma solução promissora para a classificação do sono, pois podem modelar o contexto temporal sem se restringir a padrões de curto prazo, ou depender de suposições em função do tempo absoluto, ou seja, com uma ligação explícita entre a hora da noite e a fase de sono esperada. Nomeadamente, é possível conceber as limitações que a última implica, como por exemplo, indivíduos com insónia podem ter um período de início do sono involuntariamente longo, ou em situações com interrupções do sono, casos em que as probabilidades falham [147]. Assim, ter duas camadas LSTM em paralelo aplicadas em sentido contrário, permite que a classificação de cada segmento seja condicionada pelos segmentos passados e futuros.

Apesar da vantagem comprovada deste tipo de redes, a abordagem revelou-se ineficaz neste estudo, com apenas sensibilidade para uma classe. A justificação prende-se novamente na variabilidade individual, sobretudo no padrão de sono. A rede é relativamente simples e possivelmente requer mais unidades para permitir a memorização de estruturas temporais mais complexas, e portanto, mais dados de treino. Assim, não foi possível verificar uma melhoria comparativamente aos modelos não temporais.

3.4.4 Discussão e limitações

Ao comparar os resultados obtidos com o estado da arte é possível concluir que este trabalho apresenta algumas limitações. Em primeiro, a base de dados utilizada é pequena, com uma grande variação na arquitetura do sono entre os indivíduos, que pode ser justificada pela ampla faixa etária que abrange. Assim, não foi viável considerar modelos dependentes do indivíduo ou aplicar métodos mais complexos, como em estudos recentes que demonstram uma melhoria significativa no desempenho. Em segundo, apenas é analisada a variabilidade do ECG, que pode não ser suficiente para classificar de forma completa e precisa a estrutura do sono. Por último, o sinal é propenso a artefactos, sendo que contém várias secções de ruído. Os problemas de interferência não fisiológicos podem ser parcialmente superados com o uso de filtros digitais. Contudo, artefactos de origem fisiológica como o movimento, podem ter consequências relevantes, pois a extração de intervalos RR corretos é fundamental para esta análise. Assim, pontos errados podem afetar adversamente os resultados. Alguns estudos consideram a inspeção visual do sinal, onde apenas segmentos sem ruído são selecionados [148], ou o seu pré-processamento de forma automática [149]. Portanto, este é um aspeto a melhorar, pois sobretudo os atributos no domínio do tempo podem originar *outliers*, que deterioram o desempenho da classificação.

Relativamente ao teste independente do indivíduo, os resultados da classificação de 3 classes são comparáveis ao estudo [149], com análise da VFC. Este considera 16 indivíduos, são obtidos 11 atributos, e usada uma rede neuronal com validação cruzada 10-*fold*. Foi obtida uma sensibilidade de 71%, 80% e 65% na vigília, sono NREM e REM, respetivamente, com uma exatidão de 77% e κ de 0.56. Em comparação, neste estudo foram utilizados 11 indivíduos, 7 atributos e o algoritmo RF, que permitiu atingir uma sensibilidade de 80%, 93% e 76% nas fases do sono, uma exatidão de 90% e κ de 0.71. Embora permita comparar resultados, a maioria dos estudos reporta o desempenho considerando um cenário dependente do indivíduo, que se torna mais relevante em contexto real. Alguns dos mais recentes são apresentados na tabela 3.11. A sua análise demonstra que as abordagens de *deep learning* são promissoras, sendo necessário um número considerável de dados para esta análise.

Ref.	Ano	Algoritmo	No. indivíduos	Classificação	Exatidão	κ
[150]	2018	CNN	5793	W-R-N W-R-L-D	75.3% 65.9%	0.57 0.47
[147]	2019	LSTM	292	W-R-L-D	71%	0.61
[151]	2019	LSTM	238	W-R-N W-R-L-D	84.07% 77.76%	0.58 0.55

Tabela 3.11: Estudos de classificação do sono com o eletrocardiograma.

Capítulo 4

Conclusões e trabalho futuro

4.1 Conclusões

A sonolência durante a condução apresenta-se como um problema atual e relevante que representa a causa de muitos acidentes de viação. O impacto elevado destes acidentes resulta numa motivação para o desenvolvimento de métodos que possam detetar a sonolência de forma atempada e alertar o condutor. A maioria das soluções existentes é baseada em parâmetros obtidos dos veículos ou através da monitorização do condutor, como os movimentos do volante, o piscar dos olhos ou o bocejar. Outros estudos analisam sinais fisiológicos, considerados os indicadores mais fiáveis da sonolência. Contudo, é necessário ter em consideração que estes métodos devem ser não intrusivos de modo a serem usados em contexto real de condução. Apesar da maioria das soluções oferecer resultados satisfatórios, só fazem a deteção e não a previsão da sonolência. Outro aspeto importante, é que estes são apenas sistemas reativos, ou seja, apenas emitem um alerta quando a sonolência é detetada. Assim, a inclusão de forma acessível da deteção de possíveis distúrbios do sono ou privação crónica do sono, reconhecidos como causas frequentes para o sono ao volante, constitui um passo significativo no sentido de minimizar a ocorrência destes episódios.

A utilização de *wearables* de pulso é uma tendência que prevalece. Estes dispositivos permitem a análise de múltiplos sinais de forma contínua, integrada e com baixo custo. Em particular, a frequência cardíaca é um dos sinais fisiológicos mais relevantes, tipicamente medida através de sensores PPG. Contudo, como deve ser apenas considerada em condições estacionárias, sensores ECG têm sido recentemente integrados de modo a aumentar a fiabilidade. Assim, o seu uso apresenta potencial para várias aplicações. Em primeiro, permite a deteção da sonolência do condutor. Estudos demonstram a sua eficácia, realçando a importância da capacidade de generalização. No meadamente, é necessário considerar a variabilidade interindividual, ou seja, a variação dos sinais fisiológicos entre indivíduos, um fator que é acentuado com a faixa etária. Apesar da relevância, a falta de validação da maioria das soluções existentes não promove confiança para o seu uso. Em segundo, alguns estudos evidenciam a viabilidade da previsão da sonolência de forma atempada.

Estes destacam novamente a importância de informação individual ou mesmo ambiental. Contudo, como consideram medidas maioritariamente intrusivas, seria de interesse comparar o desempenho obtido com os dispositivos em questão. Em terceiro e último, permite a classificação do sono e deteção de distúrbios. Como alternativa ao exame de polissonografia, surgiram várias soluções que reportam a arquitetura do sono de forma automática, considerando diferentes resoluções. Porém, estudos demonstram que tendem a sobrestimar a eficiência do sono. Além do mais, a sua utilidade torna-se limitada pois os dados e código que usam não são acessíveis. Por fim, no contexto da deteção de patologias, foi investigada a apneia do sono, que é caracterizada por episódios repetidos de obstrução das vias aéreas. Este diagnóstico pode ser alcançado através da análise da frequência cardíaca e saturação de oxigénio, contudo, ainda nenhum dispositivo comercial obteve validação clínica. Assim, é comprovado o valor de uma solução inovadora que conjugue estes pontos.

No âmbito deste trabalho, foi realizada a classificação automática do sono. Para este fim, foi usado um *dataset* composto por 11 indivíduos, sem nenhuma patologia de sono. Com a análise do ECG, foram aplicadas várias estratégias para comparar as abordagens mais promissoras. Foi considerada uma técnica de sobre-amostragem para lidar com dados não balanceados, e os atributos mais relevantes selecionados consoante a sua correlação. O melhor desempenho foi obtido com o algoritmo *Random Forest*, atingindo uma exatidão de 90% e 78% com κ de 0.71 e 0.64 na distinção de três e quatro classes, respetivamente. Para tal, a maior dimensão dos segmentos mostrou desempenhar um fator relevante. Além disso, manteve-se inalterado ao considerar apenas atributos no domínio do tempo, criando oportunidade para uma solução com menor custo computacional. Contudo, a abordagem revelou-se insuficiente considerando um cenário dependente do indivíduo, inclusive com modelos sequenciais, que têm em conta o contexto temporal. Os resultados podem ser justificados pela variabilidade entre indivíduos, em particular na arquitetura do sono. Assim, conclui-se que o estudo seria beneficiado com um maior número de dados.

4.2 Trabalho futuro

Como trabalho futuro, existem alguns aspetos que merecem reflexão. Em primeiro, é essencial recolher mais dados que permitam abordar a variabilidade individual, e melhorar o desempenho. Assim, seria possível testar abordagens de *deep learning*, tanto a nível da análise sequencial, como a possibilidade de obter atributos de forma automática. Em segundo, seria importante o estudo mais aprofundado do processamento do sinal, que exige um maior conhecimento do domínio. Por fim, no sentido do objetivo do projeto, avaliar o impacto da inclusão de informação adicional obtida de um dispositivo de pulso.

Referências

- [1] Ben Sawyer, Grace Teo, and Mustapha Mouloua. Driveid: Safety innovation through individualisation. *Work*, 41, 01 2012.
- [2] Muhammad Qasim Khan and Sukhan Lee. A comprehensive survey of driving monitoring and assistance systems. In *Sensors*, 2019.
- [3] Jonathan Rolison, Shirley Dorchin-Regev, Salissou Moutari, and Aidan Feeney. What are the factors that contribute to road accidents? an assessment of law enforcement views, ordinary drivers' opinions, and road accident records. *Accident Analysis Prevention*, 115, 06 2018.
- [4] Jonathan Harvy, N.v Thakor, Anastasios Bezerianos, and Junhua Li. Between-frequency topographical and dynamic high-order functional connectivity for driving drowsiness assessment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, PP:1–1, 01 2019.
- [5] Stéphanie Bioulac, Jean-Arthur Micoulaud Franchi, Mickael Arnaud, Patricia Sagaspe, Nicholas Moore, Francesco Salvo, and Pierre Philip. Risk of motor vehicle accidents related to sleepiness at the wheel: a systematic review and meta-analysis. *Sleep*, 40(10), 2017.
- [6] Pierre Thiffault and Jacques Bergeron. Monotony of road environment and driver fatigue: A simulator study. *Accident; analysis and prevention*, 35:381–91, 06 2003.
- [7] Wuhong Wang and Geert Wets. Computational intelligence for traffic and mobility. In *Atlantis Computational Intelligence Systems*, 2013.
- [8] Rosa Catarino, Jorge Spratley, Isabel Catarino, Nuno Lunet, and Manuel Pais-Clemente. Sleepiness and sleep-disordered breathing in truck drivers risk analysis of road accidents. *Sleep breathing = Schlaf Atmung*, 18, 04 2013.
- [9] Anne McCartt, John Rohrbaugh, Mark Hammer, and Sandra Fuller. Factors associated with falling asleep at the wheel among long-distance truck drivers. *Accident; analysis and prevention*, 32:493–504, 08 2000.
- [10] Rama Greenfield, Ellen Busink, Cybele C Peng Wong, Eva Fleur Riboli-Sasco, Geva Greenfield, Azeem Majeed, Josip Car, and Petra A Wark. Truck drivers' perceptions on wearable devices and health promotion: a qualitative study. In *BMC Public Health*, 2016.
- [11] Riccardo Stoohs, Christian Guilleminault, and William Dement. Sleep apnea and hypertension in commercial truck drivers. *Sleep*, 16:S11–3; discussion 13, 01 1994.

- [12] Christian Collet and Oren Musicant. Associating vehicles automation with drivers functional state assessment systems: A challenge for road safety in the future. *Frontiers in Human Neuroscience*, 13, 04 2019.
- [13] Charlotte Jacobé de Naurois, Christophe Bourdin, Anca Stratulat, Emmanuelle Diaz, and Jean-Louis Vercher. Detection and prediction of driver drowsiness using artificial neural network models. *Accident; analysis and prevention*, 126, 12 2017.
- [14] Albert Kircher, Marcus Uddman, and Jesper Sandin. *Vehicle control and drowsiness*. Statens väg-och transportforskningsinstitut, 2002.
- [15] Eleonora Tobaldini, Giorgio Costantino, Monica Solbiati, Chiara Cogliati, Tomas Kara, Lino Nobili, and Nicola Montano. Sleep, sleep deprivation, autonomic nervous system and cardiovascular diseases. *Neuroscience Biobehavioral Reviews*, 74, 07 2016.
- [16] Stuti Jaiswal, Robert Owens, and Atul Malhotra. Raising awareness about sleep disorders. *Lung India : Official Organ of Indian Chest Society*, 34:262–268, 05 2017.
- [17] Chamara Senaratna, Jennifer Perret, Caroline Lodge, Adrian Lowe, Brittany Campbell, Melanie Matheson, Garun Hamilton, and Shyamali Dharmage. Prevalence of obstructive sleep apnea in the general population: A systematic review. *Sleep Medicine Reviews*, 34, 07 2016.
- [18] Thomas Penzel, Christoph Schöbel, and Ingo Fietze. New technology to assess sleep apnea: Wearables, smartphones, and accessories. *F1000Research*, 7:413, 03 2018.
- [19] Mohamed S Ameen, Lok Man Cheung, Theresa Hauser, Michael A Hahn, and Manuel Schabus. About the accuracy and problems of consumer devices in the assessment of sleep. *Sensors*, 19(19):4160, 2019.
- [20] Faith S. Luyster, Jr. Strollo, Patrick J., Phyllis C. Zee, and James K. Walsh. Sleep: A Health Imperative. *Sleep*, 35(6):727–734, 06 2012.
- [21] Juha Kortelainen, Martin Mendez, Anna Bianchi, Matteo Matteucci, and Sergio Cerutti. Sleep staging based on signals acquired through bed sensor. *IEEE Transactions on Information Technology in Biomedicine*, 14:776–785, 05 2010.
- [22] Giuseppe Curcio, Michele Ferrara, and Luigi De Gennaro. Sleep loss, learning capacity and academic performance. *Sleep medicine reviews*, 10:323–37, 11 2006.
- [23] Siobhan Banks. Behavioral and physiological consequences of sleep restriction. *Journal of clinical sleep medicine*, 3(05):519–528, 2007.
- [24] Namni Goel, Hengyi Rao, Jeffrey Durmer, and David Dinges. Neurocognitive consequences of sleep deprivation. *Seminars in neurology*, 29:320–39, 10 2009.
- [25] Linda Ng Boyle, Jon Tippin, Amit Paul, and Matthew Rizzo. Driver performance in the moments surrounding a microsleep. *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(2):126 – 136, 2008.
- [26] Orfeu M Buxton and Enrico Marcelli. Short and long sleep are positively associated with obesity, diabetes, hypertension, and cardiovascular disease among adults in the united states. *Social science & medicine*, 71(5):1027–1036, 2010.

- [27] Olga Tkachenko and David F. Dinges. Interindividual variability in neurobehavioral response to sleep loss: A comprehensive review. *Neuroscience Biobehavioral Reviews*, 89:29–48, 2018.
- [28] Brandon Lu and Phyllis Zee. Circadian rhythm sleep disorders. *Chest*, 130:1915–23, 01 2007.
- [29] Kristin Eckel-Mahan and Paolo Sassone-Corsi. Metabolism and the circadian clock converge. *Physiological reviews*, 93(1):107–135, 2013.
- [30] Danyella Silva Pereira, Sergio Tufik, and Mario Pedrazzoli. Moléculas que marcam o tempo: implicações para os fenótipos circadianos. *Revista Brasileira de Psiquiatria*, 2009.
- [31] MH Hastings, ES Maywood, and AB Reddy. Two decades of circadian time. *Journal of neuroendocrinology*, 20(6):812–819, 2008.
- [32] Lirong Zhu and Phyllis C Zee. Circadian rhythm sleep disorders. *Neurologic clinics*, 30(4):1167–1191, 2012.
- [33] Dzmitry Aliakseyeu, Jia Du, Elly Zwartkruis-Pelgrim, and Sriram Subramanian. Exploring interaction strategies in the context of sleep. volume 6948, pages 19–36, 09 2011.
- [34] Matthew Walker. *Why we sleep: Unlocking the power of sleep and dreams*. Simon and Schuster, 2017.
- [35] Lorena Teresinha Consalter Geib, Alfredo Cataldo Neto, Ricardo Wainberg, and Magda Lahorgue Nunes. Sono e envelhecimento. *Rev Psiquiatr Rio Gd Sul*, 25(3):453–65, 2003.
- [36] Regina Maria França Fernandes. O sono normal. *Medicina (Ribeirao Preto. Online)*, 39(2):157–168, 2006.
- [37] Patrick Smallwood, Davin K. Quinn, and Theodore A. Stern. 22 - patients with disordered sleep. In Theodore A. Stern, Gregory L. Fricchione, Ned H. Cassem, Michael S. Jellinek, and Jerrold F. Rosenbaum, editors, *Massachusetts General Hospital Handbook of General Hospital Psychiatry (Sixth Edition)*, pages 289 – 302. W.B. Saunders, Saint Louis, sixth edition edition, 2010.
- [38] Shing-Tai Pan, Chih-En Kuo, Jian-Hong Zeng, and Sheng-Fu Liang. A transition-constrained discrete hidden markov model for automatic sleep staging. *Biomedical engineering online*, 11:52, 08 2012.
- [39] Aleksandar Čolić, Oge Marques, and Borko Furht. *Driver Drowsiness Detection and Measurement Methods*, pages 7–18. Springer International Publishing, Cham, 2014.
- [40] Arun Sahayadhas, Kenneth Sundaraj, and Murugappan M. Drowsiness detection during different times of day using multiple features. *Australasian physical engineering sciences in medicine / supported by the Australasian College of Physical Scientists in Medicine and the Australasian Association of Physical Sciences in Medicine*, 36, 05 2013.
- [41] Alina Mashko. Subjective methods for assessment of driver drowsiness. *Acta Polytechnica CTU Proceedings*, 12:64, 12 2017.

- [42] Thomas Kunding, Andreas Riener, and Nikoletta Sofra. A robust drowsiness detection method based on vehicle and driver vital data. *Mensch und Computer 2017-Workshopband: Spielend einfach interagieren*, 13:433, 2018.
- [43] Muhammad Awais, Nasreen Badruddin, and Micheal Driberg. A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability. *Sensors*, 17(9):1991, 2017.
- [44] Pushpa Choudhary and Nagendra R. Velaga. Analysis of vehicle-based lateral performance measures during distracted driving due to phone use. *Transportation Research Part F: Traffic Psychology and Behaviour*, 44:120 – 133, 2017.
- [45] Mkhusele Ngxande, Jules-Raymond Tapamo, and Michael Burke. Driver drowsiness detection using behavioral measures and machine learning techniques: A review of state-of-art techniques. In *2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech)*, pages 156–161. IEEE, 2017.
- [46] Arun Sahayadhas, Kenneth Sundaraj, and M. Murugappan. Detecting driver drowsiness based on sensors: A review. In *Sensors*, 2012.
- [47] Muhammad Ramzan, Hikmat Ullah Khan, Shahid Mahmood Awan, Amina Ismail, Mahwish Ilyas, and Ahsan Mahmood. A survey on state-of-the-art drowsiness detection techniques. *IEEE Access*, 7:61904–61919, 2019.
- [48] Khalid Ali I Aboalayon, Miad Faezipour, Wafaa S Almuhammadi, and Saeid Moslehpour. Sleep stage classification using eeg signal analysis: a comprehensive survey and new investigation. *Entropy*, 18(9):272, 2016.
- [49] Hamed Azami, Hamid Hassanpour, Javier Escudero, and Saeid Sanei. An intelligent approach for variable size segmentation of non-stationary signals. *Journal of Advanced Research*, 6(5):687 – 698, 2015.
- [50] U Rajendra Acharya, S Vinitha Sree, G Swapna, Roshan Joy Martis, and Jasjit S Suri. Automated eeg analysis of epilepsy: a review. *Knowledge-Based Systems*, 45:147–165, 2013.
- [51] Deepu Kurian, Johnson Joseph PL, Krishnaja Radhakrishnan, and Arun A Balakrishnan. Drowsiness detection using photoplethysmography signal. In *2014 Fourth international conference on advances in computing and communications*, pages 73–76. IEEE, 2014.
- [52] Meng Xiao, Hong Yan, Jinzhong Song, Yuzhou Yang, and Xianglin Yang. Sleep stages classification based on heart rate variability and random forest. *Biomedical Signal Processing and Control*, 8(6):624 – 633, 2013.
- [53] Koichi Fujiwara, Erika Abe, Keisuke Kamata, Chikao Nakayama, Yoko Suzuki, Toshitaka Yamakawa, Toshihiro Hiraoka, Manabu Kano, Yukiyoshi Sumi, Fumi Masuda, et al. Heart rate variability-based driver drowsiness detection and its validation with eeg. *IEEE Transactions on Biomedical Engineering*, 66(6):1769–1778, 2018.
- [54] Jinwoo Kim and Miyoung Shin. Utilizing hrv-derived respiration measures for driver drowsiness detection. *Electronics*, 8(6):669, 2019.

- [55] Farideh Ebrahimi, Seyed-Kamaleddin Setarehdan, and Homer Nazeran. Automatic sleep staging by simultaneous analysis of ecg and respiratory signals in long epochs. *Biomedical Signal Processing and Control*, 18:69 – 79, 2015.
- [56] Dwaipayan Biswas, Neide Simões-Capela, Chris Van Hoof, and Nick Van Helleputte. Heart rate estimation from wrist-worn photoplethysmography: A review. *IEEE Sensors Journal*, 19(16):6560–6570, 2019.
- [57] Noor Aimie-Salleh, Nurul Ghani, Nurhafiezah Hasanudin, and Siti Shafie. *Heart Rate Variability Recording System Using Photoplethysmography Sensor*. 11 2019.
- [58] Jasper Gielen and Jean-Marie Aerts. Feature extraction and evaluation for driver drowsiness detection based on thermoregulation. *Applied Sciences*, 9(17):3555, 2019.
- [59] Alexander Boe, Lori Koch, Megan O’Brien, Nicholas Shawen, John Rogers, Richard Lieber, Kathryn Reid, Phyllis Zee, and Arun Jayaraman. Automating sleep stage classification using wireless, wearable sensors. *npj Digital Medicine*, 2:131, 12 2019.
- [60] Thomas Kundinger, Phani Krishna Yalavarthi, Andreas Riener, Philipp Wintersberger, and Clemens Schartmüller. Feasibility of smart wearables for driver drowsiness detection and its potential among different age groups. *International Journal of Pervasive Computing and Communications*, 2020.
- [61] Thomas Kundinger, Nikoletta Sofra, and Andreas Riener. Assessment of the potential of wrist-worn wearable sensors for driver drowsiness detection. *Sensors*, 20(4):1029, 2020.
- [62] StopSleep. Anti-sleep alarm. <https://www.stopsleep.co.uk/>. Acedido a 20 de fevereiro de 2020.
- [63] Neurotone. Sleepalert. <https://neuro-tone.com/eng>. Acedido a 4 de março de 2020.
- [64] Smart Cap Tech. Life. <http://www.smartcaptech.com/life-smart-cap/>. Acedido a 4 de março de 2020.
- [65] Vigo. Vigo technologies. <https://www.wearvigo.com/shop>. Acedido a 4 de março de 2020.
- [66] Optalert. Eagle industrial. <https://www.optalert.com/explore-products/eagle-industrial/>. Acedido a 4 de março de 2020.
- [67] Boon-Leng Lee, Boon Giin Lee, Gang Li, and Wan-Young Chung. Wearable driver drowsiness detection system based on smartwatch. 11 2014.
- [68] Lee Boon Leng, Lee Boon Giin, and Wan-Young Chung. Wearable driver drowsiness detection system based on biomedical and motion sensors. In *2015 IEEE SENSORS*, pages 1–4. IEEE, 2015.
- [69] Boon-Leng Lee, Boon Giin Lee, and Wan-Young Chung. Standalone wearable driver drowsiness detection system in a smartwatch. *IEEE Sensors Journal*, 16:5444–5451, 07 2016.
- [70] M. Choi, G. Koo, M. Seo, and S. W. Kim. Wearable device-based system to monitor a driver’s stress, fatigue, and drowsiness. *IEEE Transactions on Instrumentation and Measurement*, 67(3):634–645, March 2018.

- [71] Hyeonjeong Lee, Jaewon Lee, and Miyoung Shin. Using wearable ecg/ppg sensors for driver drowsiness detection based on distinguishable pattern of recurrence plots. *Electronics*, 8:192, 02 2019.
- [72] Mohammed Misbhauddin, AlReem AlMutlaq, Alaa Almithn, Norah Alshukr, and Maryam Aleesa. Real-time driver drowsiness detection using wearable technology. In *Proceedings of the 4th International Conference on Smart City Applications*, pages 1–6, 2019.
- [73] Charlotte Jacobé de Naurois, Christophe Bourdin, Clément Bougard, and Jean-Louis Vercher. Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness. *Accident Analysis & Prevention*, 121:118–128, 2018.
- [74] Kosuke Kaida, Torbjörn ÅKERSTEDT, Göran Kecklund, Jens P Nilsson, and John Axelsson. Use of subjective and physiological indicators of sleepiness to predict performance during a vigilance task. *Industrial health*, 45(4):520–526, 2007.
- [75] Atsuo Murata, Yukio Ohta, and Makoto Moriwaka. Multinomial logistic regression model by stepwise method for predicting subjective drowsiness using performance and behavioral measures. In Ravindra Goonetilleke and Waldemar Karwowski, editors, *Advances in Physical Ergonomics and Human Factors*, pages 665–674, Cham, 2016. Springer International Publishing.
- [76] A. Watson and G. Zhou. Microsleep prediction using an ekg capable heart rate monitor. In *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pages 328–329, June 2016.
- [77] Yulan Liang, William Horrey, Mark Howard, Michael Lee, Clare Anderson, Michael Shreeve, Conor O’Brien, and Charles Czeisler. Prediction of drowsiness events in night shift workers during morning driving. *Accident Analysis & Prevention*, 126, 11 2017.
- [78] Junhua Wang, Shuaiyi Sun, Shouen Fang, Ting Fu, and Joshua Stipanovic. Predicting drowsy driving in real-time situations: using an advanced driving simulator, accelerated failure time model, and virtual location-based services. *Accident Analysis & Prevention*, 99:321–329, 2017.
- [79] Vanessa Ibáñez, Josep Silva, and Omar Cauli. A survey on sleep assessment methods. *PeerJ*, 6:e4849, 2018.
- [80] Olivia Walch, Yitong Huang, Daniel Forger, and Cathy Goldstein. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep*, 42(12), 08 2019. zsz180.
- [81] M Zambotti DE, Nicola Cellini, Aimee Goldstone, Ian M Colrain, and Fiona C Baker. Wearable sleep technology in clinical and research settings. *Medicine and science in sports and exercise*, 51(7):1538–1557, 2019.
- [82] Zachary Beattie, Y Oyang, A Statan, A Ghoreyshi, A Pantelopoulos, A Russell, and CJPM Heneghan. Estimation of sleep stages in a healthy adult population from optical plethysmography and accelerometer signals. *Physiological measurement*, 38(11):1968, 2017.
- [83] Mustafa Radha, Pedro Fonseca, Marco Ross, Andreas Cerny, Peter Anderer, and Ronald M Aarts. Lstm knowledge transfer for hrv-based sleep staging. *arXiv preprint arXiv:1809.06221*, 2018.

- [84] Xin Zhang, Weixuan Kou, Eric I-Chao Chang, He Gao, Yubo Fan, and Yan Xu. Sleep stage classification based on multi-level feature learning and recurrent neural networks via wearable device. *Computers in Biology and Medicine*, 103:71 – 81, 2018.
- [85] Pedro Fonseca, Tim Weysen, Maaïke S Goelema, Els IS Møst, Mustafa Radha, Charlotte Lunsingh Scheurleer, Leonie van den Heuvel, and Ronald M Aarts. Validation of photoplethysmography-based sleep staging compared with polysomnography in healthy middle-aged adults. *Sleep*, 40(7):zsx097, 2017.
- [86] Kwang Suk Park and Sang Ho Choi. Smart technologies toward sleep monitoring at home. *Biomedical engineering letters*, 9(1):73–85, 2019.
- [87] Tuck. How do sleep trackers work? <https://www.tuck.com/how-sleep-trackers-work/>. Acedido a 12 de março de 2020.
- [88] Matti Molkari, Mirja Tenhunen, Adrian Tarniceriu, Antti Vehkaoja, Sari-Leena Himanen, and Esa Räsänen. Non-linear heart rate variability measures in sleep stage analysis with photoplethysmography. In *2019 Computing in Cardiology (CinC)*, pages Page–1. IEEE, 2019.
- [89] I. Fedorin, K. Slyusarenko, W. Lee, and N. Sakhnenko. Sleep stages classification in a healthy people based on optical plethysmography and accelerometer signals via wearable devices. In *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, pages 1201–1204, July 2019.
- [90] Fitbit. Fitbit technology. <https://www.fitbit.com/us/technology>. Acedido a 21 de março de 2020.
- [91] Polar. Polar sleep plus. <https://www.polar.com/en/smart-coaching/polar-sleep-plus>. Acedido a 21 de março de 2020.
- [92] Bodymatter. Sleepwatch — the best sleep is made. <https://www.sleepwatchapp.com>. Acedido a 21 de março de 2020.
- [93] Sleep Cycle. Sleep cycle alarm clock. <https://www.sleepcycle.com>. Acedido a 21 de março de 2020.
- [94] SleepScore. The best sleep monitoring app. <https://www.sleepscore.com/sleepscore-app/>. Acedido a 21 de março de 2020.
- [95] Beddit. Beddit sleep monitor. <https://www.beddit.com>. Acedido a 21 de março de 2020.
- [96] Withings. Withings sleep tracking mat. <https://www.withings.com/mx/en/sleep>. Acedido a 21 de março de 2020.
- [97] Tomorrow Magazine. Tomorrow sleeptracker monitor. <https://www.tomorrowsleep.com/sleep-tracker-sleep-monitoring-system>. Acedido a 21 de março de 2020.
- [98] Dreem. Dreem 2. <https://dreem.com>. Acedido a 21 de março de 2020.
- [99] Philips. Smartsleep deep sleep headband. <https://www.usa.philips.com/c-e/smartsleep/deep-sleep-headband.html>. Acedido a 21 de março de 2020.
- [100] Sleep Shepherd. Sleep optimizer and tracker. <https://sleepshepherd.com>. Acedido a 21 de março de 2020.

- [101] SleepScore. Sleepscore max sleep improvement system. <https://www.sleepscore.com/sleepscore-max-sleep-tracker/>. Acedido a 21 de março de 2020.
- [102] ResMed. S plus by resmed sleep tracker : The smarter sleep solution. <https://splus.resmed.com>. Acedido a 21 de março de 2020.
- [103] André Pinho, Nuno Pombo, Bruno MC Silva, Kouamana Bousson, and Nuno Garcia. Towards an accurate sleep apnea detection based on ecg signal: The quintessential of a wise feature selection. *Applied Soft Computing*, 83:105568, 2019.
- [104] Philip De Chazal, Conor Heneghan, Elaine Sheridan, Richard Reilly, Philip Nolan, and Mark O'Malley. Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea. *IEEE Transactions on Biomedical Engineering*, 50(6):686–696, 2003.
- [105] Changyue Song, Kaibo Liu, Xi Zhang, Lili Chen, and Xiaochen Xian. An obstructive sleep apnea detection approach using a discriminative hidden markov model from ecg signals. *IEEE Transactions on Biomedical Engineering*, 63(7):1532–1542, 2015.
- [106] Tao Wang, Changhua Lu, Guohao Shen, and Feng Hong. Sleep apnea detection from a single-lead ecg signal with automatic feature-extraction through a modified lenet-5 convolutional neural network. *PeerJ*, 7:e7731, 2019.
- [107] Grégoire Surrel, Amir Aminifar, Francisco Rincón, Srinivasan Murali, and David Atienza. Online obstructive sleep apnea detection on medical wearable sensors. *IEEE transactions on biomedical circuits and systems*, 12(4):762–773, 2018.
- [108] Thomas Penzel, George B Moody, Roger G Mark, Ary L Goldberger, and J Hermann Peter. The apnea-ecg database. In *Computers in Cardiology 2000. Vol. 27 (Cat. 00CH37163)*, pages 255–258. IEEE, 2000.
- [109] C. W. Sia, M. Khalil-Hani, N. Shaikh-Husin, and K. H. Boon. Sleep apnea event detection system based on heart rate variability analysis. In Jemal H. Abawajy, Mohamed Othman, Rozaida Ghazali, Mustafa Mat Deris, Hairulnizam Mahdin, and Tutut Herawan, editors, *Proceedings of the International Conference on Data Engineering 2015 (DaEng-2015)*, pages 629–637, Singapore, 2019. Springer Singapore.
- [110] Kunyang Li, Weifeng Pan, Yifan Li, Qing Jiang, and Guanzheng Liu. A method to detect sleep apnea based on deep neural network and hidden markov model using single-lead ecg signal. *Neurocomputing*, 294:94 – 101, 2018.
- [111] R. K. Pathinarupothi, R. Vinaykumar, E. Rangan, E. Gopalakrishnan, and K. P. Soman. Instantaneous heart rate as a robust feature for sleep apnea severity detection using deep learning. In *2017 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, pages 293–296, 2017.
- [112] Sally K Longmore, Gough Y Lui, Ganesh Naik, Paul P Breen, Bin Jalaludin, and Gaetano D Gargiulo. A comparison of reflective photoplethysmography for detection of heart rate, blood oxygen saturation, and respiration rate at various anatomical locations. *Sensors*, 19(8):1874, 2019.
- [113] Caleb Phillips, Daniyal Liaqat, Moshe Gabel, and Eyal de Lara. Wrist02–reliable peripheral oxygen saturation readings from wrist-worn pulse oximeters. *arXiv preprint arXiv:1906.07545*, 2019.

- [114] Shu-Han Fan, Chia-Ching Chou, Wei-Chen Chen, and Wai-Chi Fang. Real-time obstructive sleep apnea detection from frequency analysis of edr and hrv using lomb periodogram. In *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pages 5989–5992. IEEE, 2015.
- [115] Hemant Sharma and KK Sharma. An algorithm for sleep apnea detection from single-lead ecg using hermite basis functions. *Computers in biology and medicine*, 77:116–124, 2016.
- [116] André Miguel da Silva Pinho, Nuno Pombo, and Nuno M Garcia. Sleep apnea detection using a feed-forward neural network on ecg signal. In *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pages 1–6. IEEE, 2016.
- [117] Xiaolong Liang, Xing Qiao, and Yongtao Li. Obstructive sleep apnea detection using combination of cnn and lstm techniques. In *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pages 1733–1736. IEEE, 2019.
- [118] G. B. Papini, P. Fonseca, J. Margarito, M. M. van Gilst, S. Overeem, J. W. M. Bergmans, and R. Vullings. On the generalizability of ecg-based obstructive sleep apnea monitoring: merits and limitations of the apnea-ecg database. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6022–6025, 2018.
- [119] Withings. Discover scanwatch, the newest hybrid smartwatch from withings. <https://www.withings.com/de/en/scanwatch>. Acedido a 05 de abril de 2020.
- [120] Mario Giovanni Terzano, Liborio Parrino, Arianna Smerieri, Ronald Chervin, Sudhansu Chokroverty, Christian Guilleminault, Max Hirshkowitz, Mark Mahowald, Harvey Moldofsky, Agostino Rosa, et al. Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (cap) in human sleep. *Sleep medicine*, 3(2):187–199, 2002.
- [121] Bob Kemp, Aeilko H Zwinderman, Bert Tuk, Hilbert AC Kamphuisen, and Josefien JL Obery. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the eeg. *IEEE Transactions on Biomedical Engineering*, 47(9):1185–1194, 2000.
- [122] Mohammad Rezaei. EEG/EOG/EMG data from a cross sectional study on psychophysiological insomnia and normal sleep subjects, 09 2017.
- [123] Vincent van Hees, Sarah Charman, and Kirstie Anderson. Newcastle polysomnography and accelerometer data, Jan 2018.
- [124] Olivia Walch. Motion and heart rate from a wrist-worn wearable and labeled sleep from polysomnography (version 1.0.0). PhysioNet, 2019.
- [125] Empatica. E4 wristband. <https://www.empatica.com/en-eu/research/e4/>. Acedido a 5 de março de 2020.
- [126] Bitalino. Bitalino. <https://bitalino.com/en/>. Acedido a 5 de março de 2020.
- [127] Libelium. Mysignals. <http://www.my-signals.com>. Acedido a 5 de março de 2020.
- [128] Biovotion. Everion. <https://www.biovotion.com/everion/>. Acedido a 7 de março de 2020.
- [129] Polar. Heart rate sensors. <https://www.polar.com/en/products/heart-rate-sensors>. Acedido a 7 de março de 2020.

- [130] PulseOn. Ohr tracker. <https://pulseon.com/tech/ohr-tracker>. Acedido a 19 de março de 2020.
- [131] Liliana Barrios, Pietro Oldrati, Silvia Santini, and Andreas Lutterotti. Evaluating the accuracy of heart rate sensors based on photoplethysmography for in-the-wild analysis. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 251–261, 2019.
- [132] Julia Pietilä, Saeed Mehrang, Johanna Tolonen, Elina Helander, Holly Jimison, Misha Pavel, and Ilkka Korhonen. Evaluation of the accuracy and reliability for photoplethysmography based heart rate and beat-to-beat detection during daily activities. In *EMBECC & NBC 2017*, pages 145–148. Springer, 2017.
- [133] Pedro Gomes, Petra Margaritoff, and Hugo Silva. pyHRV: Development and evaluation of an open-source python toolbox for heart rate variability (HRV). In *Proc. Int’l Conf. on Electrical, Electronic and Computing Engineering (IcETRAN)*, pages 822–828, 2019.
- [134] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [135] François Chollet et al. Keras. <https://keras.io>, 2015.
- [136] Hamid Baghaee, Dragan Mlakić, Srete Nikolovski, and Tomislav Dragicevic. Support vector machine-based islanding and grid fault detection in active distribution networks. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, PP:1–19, 05 2019.
- [137] Medium. Story telling for linear discriminant analysis(lda). <https://medium.com/analytics-vidhya/story-telling-for-linear-discriminant-analysis-lda-a53261afc2a9>. Acedido a 19 de abril de 2020.
- [138] Ruchi Gajjar and Tanish Hemalbhai Zaveri. Defocus blur radius classification using random forest classifier. *2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC)*, pages 219–223, 2017.
- [139] Saleh Alaliyat. Video-based fall detection in elderly’s houses. Master’s thesis, 2008.
- [140] Christopher Olah. Understanding lstm networks. <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Acedido a 20 de maio de 2020.
- [141] Medium. Unbalanced datasets & what to do. <https://medium.com/strands-tech-corner/unbalanced-datasets-what-to-do-144e0552d9cd>. Acedido a 11 de maio de 2020.
- [142] Yonghua Xie, Yurong Liu, and Qingqiu Fu. Imbalanced data sets classification based on svm for sand-dust storm warning. *Discrete Dynamics in Nature and Society*, 2015:1–8, 10 2015.
- [143] Hien M Nguyen, Eric W Cooper, and Katsuari Kamei. Borderline over-sampling for imbalanced data classification. *International Journal of Knowledge Engineering and Soft Data Paradigms*, 3(1):4–21, 2011.
- [144] Girish Chandrashekar and Ferat Sahin. A survey on feature selection methods. *Computers Electrical Engineering*, 40(1):16 – 28, 2014. 40th-year commemorative issue.

- [145] Seong Park, Jin Mo Goo, and Chan-Hee Jo. Receiver operating characteristic (roc) curve: Practical review for radiologists. *Korean journal of radiology : official journal of the Korean Radiological Society*, 5:11–8, 03 2004.
- [146] Rok Blagus and Lara Lusa. Smote for high-dimensional class-imbalanced data. *BMC bioinformatics*, 14:106, 03 2013.
- [147] Mustafa Radha, Pedro Fonseca, Arnaud Moreau, Marco Ross, Andreas Cerny, Peter Anderer, Xi Long, and Ronald M Aarts. Sleep stage classification from heart-rate variability using long short-term memory neural networks. *Scientific reports*, 9(1):1–11, 2019.
- [148] Bülent Yılmaz, Musa H Asyalı, Eren Arıkan, Sinan Yetkin, and Fuat Özgen. Sleep stage and obstructive apneic epoch classification using single-lead ecg. *Biomedical engineering online*, 9(1):39, 2010.
- [149] Ran Wei, Xinghua Zhang, Jinhai Wang, and Xin Dang. The research of sleep staging based on single-lead electrocardiogram and deep neural network. *Biomedical engineering letters*, 8(1):87–93, 2018.
- [150] Qiao Li, Qichen Li, Chengyu Liu, Supreeth P Shashikumar, Shamim Nemati, and Gari D Clifford. Deep learning in the cross-time frequency domain for sleep staging from a single-lead electrocardiogram. *Physiological measurement*, 39(12):124005, 2018.
- [151] Yuhui Wei, Xia Qi, Huaning Wang, Zhian Liu, Gang Wang, and Xiangguo Yan. A multi-class automatic sleep staging method based on long short-term memory network using single-lead electrocardiogram signals. *IEEE Access*, 7:85959–85970, 2019.