Journal of Health Informatics

J Health Inform. 2026;18:e1838
DOI: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1838

ARTIGO COMPLETO
CBIS 2026

Triagem da apneia obstrutiva do sono por meio de inteligência artificial aplicada à imagem da língua e a variáveis clínicas

Screening for obstructive sleep apnea using artificial intelligence applied to tongue images and clinical variables

Afonso Ueslei da Fonseca[1] ✉, Lívia Pinheiro Siqueira[2], Juliana Paula Felix[1], Marcelo Fouad Rabahi[2], Fabrizzio Soares[1]

ORCID: 0000-0001-5517-2051 · 0000-0002-7647-5534 · 0000-0003-4095-1639 · 0000-0002-4050-5906 · 0000-0003-1598-1377

[1] Instituto de Informática, Universidade Federal de Goiás, Brasil.
[2] Faculdade de Medicina, Universidade Federal de Goiás, Brasil.

FICHA DO ARTIGO

TIPO
Artigo completo · CBIS 2026

IDENTIFICADOR
e1838

DOI
10.59681/2175-4411.v18.iEspecial.2026.1838

RECEBIDO
17/05/2026

APROVADO
20/07/2026

PROCESSO DE AVALIAÇÃO
Double Blind Review

EDITOR RESPONSÁVEL
Juliano de Souza Gaspar

ACESSO
Aberto, CC BY 4.0

AUTOR CORRESPONDENTE
Afonso Ueslei da Fonseca
afonsoueslei@ufg.br

LICENÇA

Este é um artigo publicado em acesso aberto sob a licença Creative Commons CC BY 4.0, que permite uso, distribuição e reprodução em qualquer meio, desde que o trabalho original seja corretamente citado.

CC BY 4.0

RESUMO

A apneia obstrutiva do sono é uma condição de alta prevalência e frequentemente subdiagnosticada, associada a importantes riscos à saúde, evidenciando a necessidade de métodos alternativos de rastreio acessíveis e não invasivos. **Objetivo:** Este estudo propõe um modelo de inteligência artificial para o rastreio da apneia obstrutiva do sono, combinando variáveis clínicas e descritores de imagem da língua. **Métodos:** Foi realizado um estudo observacional e transversal com 155 pacientes submetidos à polissonografia tipo I. Dados clínicos e descritores texturais, extraídos de imagens da língua segmentadas por uma rede YOLO26s, foram selecionados por ANOVA e combinados em um vetor de características utilizado como entrada para um classificador SVC-RBF treinado com validação cruzada estratificada. **Resultados:** O modelo apresentou sensibilidade de 91,43%, acurácia de 89,13% e AUC de 85,82%. **Conclusão:** A integração de dados clínicos e imagens da língua mostrou potencial como ferramenta complementar para o rastreio da apneia obstrutiva do sono.

Descritores: apneia obstrutiva do sono; polissonografia; inteligência artificial.

ABSTRACT

Obstructive sleep apnea is a highly prevalent and frequently underdiagnosed condition associated with significant health risks, highlighting the need for accessible and non-invasive alternative screening methods. **Objective:** This study proposes an artificial intelligence model for obstructive sleep apnea screening, combining clinical variables and tongue image descriptors. **Methods:** A cross-sectional observational study was conducted with 155 patients who underwent type I polysomnography. Clinical data and textural descriptors, extracted from tongue images segmented by a YOLO26s network, were selected by ANOVA and combined into a feature vector used as input for an SVC-RBF classifier trained with stratified cross-validation. **Results:** The model showed a sensitivity of 91.43%, accuracy of 89.13%, and AUC of 85.82%. **Conclusion:** The integration of clinical data and tongue images has shown potential as a complementary tool for obstructive sleep apnea screening.

Keywords: obstructive sleep apnea; polysomnography; artificial intelligence.

CBIS 2026 XXI CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE

Artigo apresentado no XXI Congresso Brasileiro de Informática em Saúde (CBIS 2026), realizado de 23 a 25 de setembro de 2026, em Brasília (DF), e publicado na seção CBIS 2026 do Journal of Health Informatics.

COMO CITAR ESTE ARTIGO: da Fonseca AU, Siqueira LP, Felix JP, Rabahi MF, Soares F. Triagem da apneia obstrutiva do sono por meio de inteligência artificial aplicada à imagem da língua e a variáveis clínicas. J Health Inform 2026;18:e1838. doi: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1838

1 INTRODUÇÃO

A apneia obstrutiva do sono (AOS) é um distúrbio respiratório caracterizado por episódios recorrentes de obstrução parcial ou completa das vias aéreas superiores (VAS) durante o sono, associados a dessaturações intermitentes, despertares recorrentes e fragmentação do sono [1,2]. Suas repercussões clínicas envolvem alterações neurocognitivas, metabólicas e cardiovasculares, com impacto sobre a qualidade de vida e elevada utilização dos serviços de saúde [3,4].

Estima-se que, globalmente, quase 1 bilhão de adultos entre 30 e 69 anos apresentam AOS em algum grau de severidade, sendo que aproximadamente 425 milhões têm as formas moderadas a graves [5]. Apesar dessa elevada prevalência, uma parcela significativa dos casos permanece sem reconhecimento clínico, evidenciando a necessidade de estratégias que aprimorem a identificação inicial de indivíduos com maior probabilidade do distúrbio [6,7].

A polissonografia (PSG) tipo I é o exame padrão-ouro para o diagnóstico da AOS [2,6]. Entretanto, seu custo, disponibilidade limitada, necessidade de estrutura especializada e tempo prolongado dificultam o acesso amplo e oportuno à confirmação diagnóstica [2,8]. Esse cenário torna relevante o desenvolvimento de métodos complementares de triagem que auxiliem a priorização para investigação formal, sem substituir os exames objetivos do sono [6,9].

Na prática clínica, instrumentos como STOP-Bang, questionário de Berlim e escala de sonolência de Epworth (ESE) são amplamente utilizados, embora apresentem desempenho variável e baixa especificidade em diferentes populações [8,10]. A combinação de marcadores clínicos, antropométricos e anatômicos pode oferecer estratégia mais robusta para a estratificação inicial de risco [10,11]. Nesse contexto, a circunferência do pescoço (CP) é a variável antropométrica mais fortemente associada à AOS [11, 12], e a língua, em razão de sua participação na patência das VAS, é reconhecida como estrutura central na fisiopatologia da doença [13,14]. A presença de marcas dentárias nas bordas laterais da língua, denominada língua crenada (LC), pode refletir desproporção entre o volume lingual e a arcada dentária e tem sido descrita como achado clínico potencialmente associado a distúrbios respiratórios do sono [15].

Com o avanço dos métodos computacionais, modelos de aprendizado de máquina (*machine learning*, ML) vêm sendo investigados como estratégias de apoio à triagem, predição de risco e tomada de decisão em saúde [9,16]. No contexto da AOS, esses modelos podem integrar variáveis clínicas, antropométricas, sintomas e características extraídas de imagens, com o objetivo de estimar a probabilidade da doença e auxiliar na triagem e priorização de pacientes para exames confirmatórios, geralmente associados a maior complexidade ou custo [9,16,17]. A integração entre dados clínicos e descritores objetivos extraídos de fotografias padronizadas da língua representa uma possibilidade ainda pouco explorada na triagem da AOS [9].

Diante desse contexto, este estudo teve como objetivo desenvolver e avaliar um modelo preditivo, com uso de inteligência artificial (IA), para triagem da AOS na tarefa binária controle versus caso, a partir da integração entre variáveis clínicas e descritores texturais de imagens da língua.

2 MÉTODOS

A Figura 1 ilustra, de forma integrada e resumida, o fluxo do método adotado: coleta de dados clínicos e de fotografias da língua em consultório; pré-processamento das variáveis e das imagens; segmentação supervisionada das imagens com rede YOLO26 [18]; extração de descritores de textura pelo método Binary Pattern of Phase Congruency (BPPC) com redimensionamento para 256×256 pixels; seleção de características por análise de variância (ANOVA); classificação binária (controle versus caso) por *Support Vector Classifier* com *kernel* radial (SVC-RBF); e avaliação por métricas de desempenho, matriz de confusão e curva ROC. As subseções seguintes detalham cada etapa.

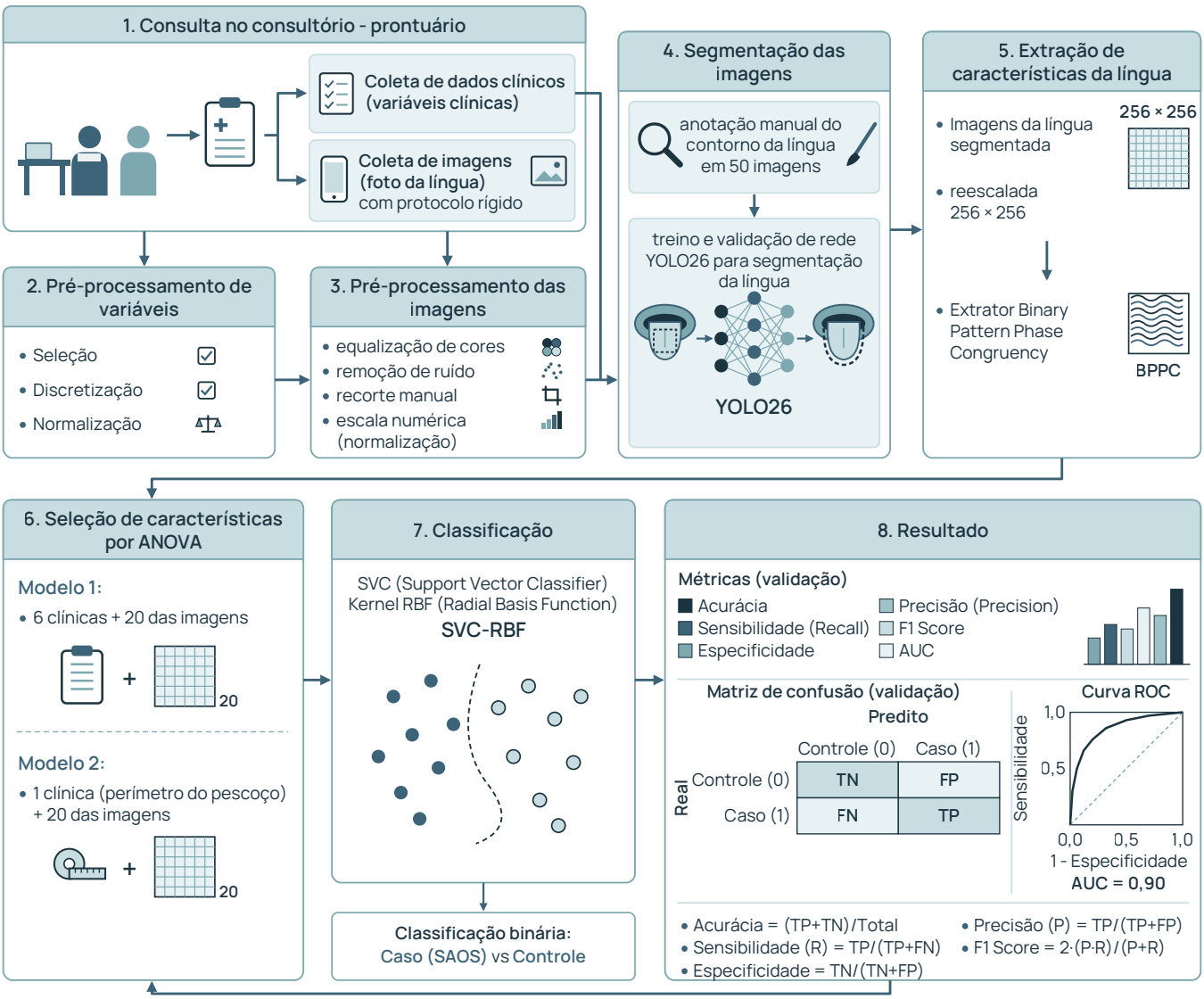
2.1 Delineamento e aspectos éticos

Trata-se de estudo observacional e transversal, com desenvolvimento de modelo preditivo para triagem da AOS. O projeto foi aprovado pelo Comitê de Ética em Pesquisa do Hospital Estadual Geral de Goiânia Dr. Alberto Rassi, sob parecer nº 7.411.443 e CAAE nº 85899425.0.0000.0035. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido (TCLE).

2.2 Participantes

A população do estudo foi composta por pacientes adultos submetidos à PSG tipo I em uma clínica especializada em medicina do sono (Clínica SleepLab, Rio de Janeiro, RJ, Brasil) nos meses de março e abril de 2025. Foram incluídos indivíduos com idade ≥ 18 anos, que apresentavam dados clínicos essenciais completos, capacidade de protrusão lingual e que assinaram o TCLE. Foram excluídos participantes com histórico de tratamento prévio para AOS, dentição incompleta, cirurgia oral recente, presença de lesões ulceradas extensas, impossibilidade de mensuração da CP, alterações ou malformações craniofaciais relevantes, bem como aqueles cujas fotografias apresentavam qualidade insuficiente para análise. A amostra final foi composta por 155 participantes, totalizando 307 imagens, com média de duas imagens por indivíduo (capturadas dos

Figura 1. Fluxograma do modelo de triagem da AOS, da coleta à classificação binária por SVC-RBF.



Fonte: Elaborado pelos autores (2026).

Tabela 1. Caracterização da amostra (n=155)

Variável	População (n=155)	Caso (n=138)	Controle (n=17)	p
Sexo masculino, n (%)	91 (58,7)	87 (63,0)	4 (23,5)	0,002*
Idade (anos)	48,5 ± 14,5	49,9 [22-80]	36,9 [18-75]	0,004*
IMC (kg/m^2)	31,1 ± 6,4	31,9 [17,0-51,4]	26,0 [20,4-33,6]	‹0,001*
CP (cm)	39,7 ± 4,2	40,4 [30-49]	34,2 [28-39]	‹0,001*

n: número de amostras; σ: desvio-padrão; *p ≤ 0,05 nível de significância. As variáveis contínuas foram comparadas pelo teste t de Student ou Mann-Whitney e a variável categórica pelo teste qui-quadrado ou Fisher.

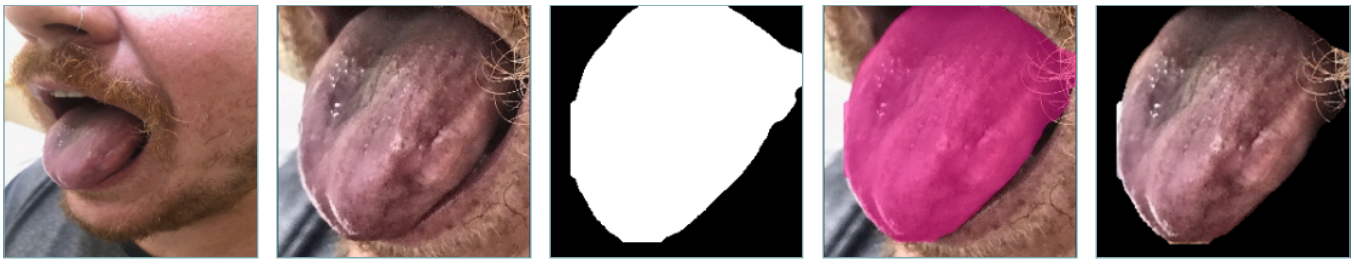
Fonte: Elaborado pelos autores (2026).

lados esquerdo e direito da língua). A Tabela 1 traz o perfil dos participantes, com média de idade de 48,5 ± 14,5 anos, predomínio do sexo masculino (58,7%) e índice de massa corporal (IMC) médio de 31,1 ± 6,4 kg/m^2. A CP média foi de 39,7 ± 4,2 cm. Na comparação entre controles e casos, os participantes com AOS (n=138) apresentaram maior idade (49,9 ± 13,8 vs. 36,9 ± 15,7 anos; p = 0,004; g = 0,92), maior proporção do sexo masculino (63,0% vs. 23,5%; p = 0,002; OR=5,54; IC95%: 1,72–17,91), maior IMC (31,9 ± 6,4 vs. 26,0 ± 3,03 kg/m^2; p < 0,001; rrb = 0,60) e maior CP (40,4 ± 3,9 vs. 34,2 ± 2,7 cm; p < 0,001; g = 1,64).

2.3 Aquisição fotográfica e variáveis clínicas

As fotografias foram capturadas por meio de *smartphones* (iPhone 11 e 12 Pro Max), seguindo um protocolo padronizado: participantes em posição sentada, postura neutra, com a língua projetada

Figura 2. Exemplo do processo de segmentação pela rede YOLO26s



Fonte: Elaborado pelos autores (2026).

de forma relaxada sobre o lábio inferior, a uma distância de 25–30 cm do dispositivo. O enquadramento restringiu-se ao terço inferior da face, com o objetivo de minimizar o risco de identificação facial.

Na etapa de pré-processamento das variáveis (passo 2), foram selecionadas para este estudo as variáveis demográficas (idade e sexo), antropométricas (peso, altura, IMC e CP) e clínicas (ESE e presença de ronco), sendo todas tratadas, neste estudo, sob a denominação geral de variáveis clínicas. Nessa mesma etapa, variáveis que não integrariam as análises, como data da coleta, nome do paciente, motivo da PSG, índice de apneia-hipopneia (IAH) e tempo com saturação de oxigênio < 90% (T90), foram removidas do conjunto de dados.

2.4 Padrão de referência

Os laudos da PSG tipo I foram elaborados conforme as diretrizes da American Academy of Sleep Medicine (AASM) [19]. O diagnóstico de AOS foi definido pelo IAH. Para a análise binária principal, definiu-se o grupo controle como participantes sem AOS (IAH<5 eventos/hora) e o grupo caso como participantes com AOS (IAH≥5 eventos/hora).

2.5 Pré-processamento e segmentação da língua

As imagens foram organizadas em duas categorias morfológicas, crenada (n=157) e não crenada (n=150), totalizando 307 fotografias. Desse total, cerca de 16% (n=50, 25 crenadas e 25 não crenadas) foram manualmente anotadas na plataforma MakeSense (ferramenta *online* gratuita para rotulagem de imagens, disponível em https://www.makesense.ai/), gerando máscaras de referência (*ground truth*) utilizadas no treinamento supervisionado de um modelo de segmentação por instância baseado na arquitetura YOLO26s-seg (variante *small*, 260 camadas, pré-treinada e ajustada ao domínio).

Todas as imagens passaram por pré-processamento (equalização de cores, redução de ruído, normalização e discretização) para minimizar variações de captura. Como entrada, utilizaram-se imagens redimensionadas para 320 × 320 pixels. O modelo foi treinado por 100 épocas, número definido para garantir a convergência dos pesos durante o processo de ajuste fino (*fine-tuning*) e a estabilização da função de perda, favorecendo a adaptação da rede às características morfológicas específicas das imagens de língua sem evidências de sobreajuste. O desempenho foi avaliado por meio das métricas AP@50, AP@50–95, IoU e DSC (mAP@50, *mean Average Precision*: precisão média considerando como verdadeiros positivos os casos em que a interseção sobre união entre a predição e a anotação de referência é maior ou igual a 0,5; mAP@50–95: precisão média calculada em múltiplos limiares de IoU, tipicamente de 0,5 a 0,95 em incrementos de 0,05, considerando diferentes níveis de sobreposição entre predição e verdade de referência; IoU, *Intersection over Union*: medida de sobreposição entre duas regiões, definida como a razão entre a área de interseção e a área de união entre a predição e a anotação de referência; DSC, *Dice Similarity Coefficient*: coeficiente de similaridade que mede o grau de sobreposição entre duas regiões, definido como duas vezes a área de interseção dividida pela soma das áreas das regiões comparadas), tanto para detecção (caixas delimitadoras) quanto para segmentação (máscaras poligonais), com o objetivo de isolar a região da língua (*region of interest*, ROI) e reduzir interferências de estruturas adjacentes nos processos subsequentes do método.

A Figura 2 ilustra as etapas do processo: (a) imagem original; (b) imagem pré-processada com a região de interesse detectada; (c) máscara segmentada da língua; (d) sobreposição da máscara sobre a imagem original; e (e) segmentação final obtida pelo modelo YOLO26s.

2.6 Extração de características de imagem

A partir da região de interesse delimitada pela máscara de segmentação, procedeu-se à extração de características de textura da língua por meio do método *BPPC* [20]. A imagem foi convertida

Tabela 2. Desempenho dos modelos SVC-RBF na classificação binária Controle vs Caso

Métrica (%)	SVC-RBF + VC	SVC-RBF + CP
Sensibilidade	91,43 ± 8,11	87,73 ± 7,49
Especificidade	70,00 ± 34,94	70,00 ± 34,96
Acurácia global	89,13 ± 8,50	85,88 ± 7,08
Acurácia balanceada	80,71 ± 18,91	78,87 ± 17,76
F1-score macro	77,00 ± 18,46	71,29 ± 13,17
AUC ROC	85,82 ± 15,72	87,20 ± 15,18

Valores em média ± desvio-padrão (validação cruzada estratificada em 10 partições). VC: variáveis clínicas (idade, sexo, IMC, CP, indicador de sonolência derivado da ESE e ronco); CP: circunferência do pescoço.

Fonte: Elaborado pelos autores (2026).

Tabela 3. Matrizes de confusão agregadas

	(a) SVC-RBF + VC		(b) SVC-RBF + CP	
	Predito controle	Predito caso	Predito controle	Predito caso
Real controle	12 (VN)	5 (FP)	12 (VN)	5 (FP)
Real caso	12 (FN)	126 (VP)	17 (FN)	121 (VP)

VN: verdadeiro negativo; FP: falso positivo; FN: falso negativo; VP: verdadeiro positivo.

Fonte: Elaborado pelos autores (2026).

para escala de cinza, redimensionada para 256 × 256 *pixels* e normalizada (escala 0–1). O BPPC foi aplicado em quatro fases distintas (0°, 90°, 180° e 270°), gerando histogramas com 177 *bins* por fase. A concatenação dos histogramas resultou em um vetor de 708 características por imagem.

2.7 Modelagem e validação

O conjunto preditivo combinou seis variáveis clínicas (idade, sexo, IMC, CP, indicador de sonolência diurna derivado da ESE e presença de ronco), previamente selecionadas por ANOVA, aos descritores BPPC. A ANOVA também foi aplicada a esses descritores de imagem, mantendo-se os 20 mais discriminativos, totalizando 26 variáveis (6 clínicas e 20 de imagem). A padronização (*StandardScaler*) foi ajustada exclusivamente nos dados de treinamento de cada *fold*.

O classificador adotado foi o SVC-RBF, com hiperparâmetros padrão do scikit-learn (C = 1,0; γ = "scale"; random_state = 42 e class_weight = "balanced") para mitigar o desbalanceamento entre as classes, sem geração sintética de amostras. A capacidade de generalização foi avaliada por validação cruzada estratificada em 10 partições (*StratifiedKFold*, shuffle=True), com o participante como unidade analítica. As análises foram realizadas em Python com scikit-learn 1.8.0, NumPy 2.4.2 e pandas 2.3.3. Para contextualização do desempenho, foi avaliada uma configuração comparativa empregando a CP associada aos descritores BPPC (SVC-RBF + CP), com o objetivo de estimar o ganho incremental da combinação completa de variáveis clínicas em relação a esse marcador antropométrico isolado, identificado pela análise ANOVA como o de maior poder discriminatório para AOS.

3 RESULTADOS

O modelo de segmentação, embora não constituísse o objetivo principal e integrasse o *pipeline* apenas como etapa de pré-processamento, foi também avaliado para verificar se a automatização não incorria em perda de informação relevante. O modelo treinado apresentou mAP@50 de 99,50%, mAP@50–95 de 92,28%, IoU de 85,76% e coeficiente de Dice de 92,26%, evidenciando elevada concordância na delimitação anatômica da língua e fornecendo suporte robusto à etapa subsequente de classificação.

Na tarefa alvo, classificação binária controle versus caso, o modelo SVC-RBF combinando variáveis clínicas e descritores BPPC (SVC-RBF + VC) apresentou sensibilidade média de 91,43 ± 8,11%, acurácia global de 89,13 ± 8,50%, acurácia balanceada de 80,71 ± 18,91%, *F1-score* macro de 77,00 ± 18,46% e AUC de 85,82 ± 15,72%. A especificidade média foi de 70,00 ± 34,94%, com elevada variabilidade entre as partições, que pode estar atribuída ao reduzido número de controles. A configuração comparativa empregando CP associada aos descritores BPPC (SVC-RBF + CP) apresentou desempenho inferior em sensibilidade, acurácia global, acurácia balanceada e *F1-score* macro, ainda que com estabilidade da especificidade e AUC ligeiramente superior numericamente (Tabela 2).

As matrizes de confusão agregadas das dez partições são apresentadas na Tabela 3. Para o modelo SVC-RBF + VC, Tabela 3(a), observaram-se 126 verdadeiros positivos, 12 verdadeiros negativos, 5 falsos positivos e

12 falsos negativos. A partir dessa matriz, o valor preditivo positivo (VPP) foi de 96,18% e o valor preditivo negativo (VPN) de 50,00%; as razões de verossimilhança positiva e negativa foram, respectivamente, 3,10 e 0,12, indicando aumento moderado da probabilidade pós-teste diante de resultado positivo e redução relevante da probabilidade de AOS diante de resultado negativo.

4 DISCUSSÃO

A investigação da AOS permanece um desafio clínico, dado que a triagem baseada apenas em instrumentos clínicos subjetivos apresenta baixa precisão e a PSG tipo I é onerosa e de acesso limitado [2,6]. Instrumentos como a ESE têm utilidade auxiliar, mas acurácia insuficiente para predição diagnóstica independente; marcadores físicos objetivos, como CP, IMC e alterações anatômicas das VAS, embora de fácil obtenção, são igualmente insuficientes quando analisados isoladamente [11,12]. Estratégias de triagem baseadas na combinação de múltiplos achados, suportadas por IA, têm-se mostrado promissoras nesse contexto [9,16].

Neste estudo, o classificador SVC-RBF + VC, integrando seis variáveis clínicas a 20 descritores BPPC extraídos de fotografias da língua, demonstrou desempenho favorável na tarefa binária controle versus caso. A sensibilidade de 91,43% e o VPN de 50,00%, associados à razão de verossimilhança negativa de 0,12, indicam que o modelo é capaz de reduzir substancialmente a probabilidade pós-teste de AOS diante de resultado negativo, característica essencial para um instrumento de triagem cujo objetivo principal é minimizar falsos negativos e identificar pacientes que demandam investigação diagnóstica formal.

A análise comparativa entre as configurações reforça a relevância da integração de variáveis. O modelo SVC-RBF + VC superou o SVC-RBF + CP em sensibilidade, acurácia global, acurácia balanceada e *F1-score* macro. Embora a configuração baseada em CP associada aos descritores tenha apresentado mesma especificidade e AUC ligeiramente superior numericamente, essa métrica não deve ser interpretada isoladamente em cenários de triagem, nos quais sensibilidade e acurácia balanceada são mais relevantes para a identificação confiável de pacientes em risco [16].

É importante ressaltar que o desempenho do modelo SVC-RBF + CP serviu como uma referência (*baseline*) antropométrica. Embora o SVC-RBF + CP tenha apresentado métricas inferiores ao complemento do erro da classe majoritária, a superioridade do modelo SVC-RBF + VC demonstra que a integração de dados clínicos amplos (idade, sexo, IMC, ronco) com descritores de imagem aporta informações preditivas que transcendem o marcador antropométrico isolado. Em comparação com o estudo de Mencar et al. [16], que utilizou validação cruzada estratificada em 10 partições para predição de gravidade, nosso modelo demonstrou competitividade ao integrar um marcador anatômico de imagem adicional, superando a acurácia de modelos baseados exclusivamente em variáveis clínicas.

Os resultados encontram convergência com a literatura. Mencar et al. [16] avaliaram diferentes algoritmos para predizer a gravidade da AOS a partir de variáveis clínicas e funcionais, observando desempenho superior de SVM e Random Forest e adotando validação cruzada estratificada em 10 partições, estratégia também empregada no presente estudo para reduzir viés otimista e estimar a capacidade de generalização. Diferentemente daquele trabalho, predominantemente baseado em variáveis clínicas, esta pesquisa incorporou descritores de imagem da língua, ampliando a abordagem por meio de um marcador anatômico adicional. Tsuiki et al. [21] demonstraram que algoritmos de ML aplicados a imagens craniofaciais podem identificar AOS grave com AUC de até 92%; contudo, dependem de equipamento radiológico especializado. O uso de fotografias da língua obtidas com *smartphone*, como no presente estudo, amplia a aplicabilidade clínica e a viabilidade da estratégia.

A elevada variabilidade das métricas (σ), especialmente da especificidade, reflete o reduzido número de controles entre as partições, típico de cenários com alta probabilidade pré-teste em clínicas do sono. Embora algumas partições tenham apresentado desempenho muito elevado, inclusive AUC=1,00, esses resultados devem ser interpretados com cautela, pois podem indicar instabilidade estatística e composição favorável das partições. Assim, a análise conjunta das métricas, da variabilidade entre partições e da matriz agregada deve ser priorizada, por proporcionar interpretação mais transparente. Do ponto de vista de aplicabilidade clínica, o modelo deve ser compreendido como uma ferramenta complementar de triagem, especialmente em ambientes nos quais a PSG tipo I é limitada e a priorização de encaminhamentos é necessária [22,23]. Um resultado negativo não deve se sobrepor ao julgamento clínico em pacientes muito sintomáticos, com comorbidades relevantes ou sinais anatômicos importantes [24].

Este estudo apresenta limitações. A amostra foi proveniente de clínica especializada em medicina do sono, com elevada prevalência de AOS, o que restringe a extrapolação dos resultados para outros cenários. O número reduzido de participantes sem AOS comprometeu a estabilidade das estimativas da classe controle, sobretudo da especificidade. Embora tenham sido empregadas técnicas para mitigar o desbalanceamento das classes, a predominância da classe positiva pode ter favorecido a alta sensibilidade observada, e a acurácia média (89,13%) foi apenas ligeiramente superior ao desempenho esperado pela classificação da classe majoritária (89,03%). Além disso, foi realizada

apenas validação interna, sem validação temporal ou multicêntrica externa, não foi empregada *nested cross-validation* [25] para ajuste formal de hiperparâmetros nem análise sistemática de explicabilidade e o método baseia-se em fotografia estática da língua em vigília, captando apenas parte do fenótipo anatômico-funcional da AOS. Portanto, os resultados devem ser interpretados como prova de conceito, requerendo validação em amostras independentes antes de sua aplicação clínica mais ampla.

A aplicação da seleção de características via ANOVA de forma separada para descritores clínicos e de imagem foi adotada para equilibrar a representatividade das modalidades no conjunto final. Contudo, reconhecemos que essa abordagem pode omitir interações sinérgicas entre as variáveis. Futuros estudos devem avaliar técnicas de seleção de características integradas, como o *Recursive Feature Elimination* (RFE), que consideram o conjunto completo de variáveis simultaneamente.

5 CONCLUSÕES

Os resultados sustentam a viabilidade de uma abordagem complementar de triagem para AOS baseada na integração entre variáveis clínicas e descritores de imagem da língua extraídos por método automatizado. Na tarefa binária controle versus caso, o classificador SVC-RBF apresentou sensibilidade de 91,43%, acurácia global de 89,13% e AUC de 85,82% na validação interna, com desempenho superior àquele observado quando a circunferência do pescoço foi utilizada como única variável clínica associada aos descritores de imagem. O modelo deve ser interpretado como prova de conceito, com potencial para auxiliar a priorização de encaminhamentos para PSG tipo I. Estudos futuros devem incluir validação externa em amostras independentes e multicêntricas, análise em cenários de menor prevalência e avaliação do impacto clínico no fluxo assistencial.

Contribuição dos autores. Os autores contribuíram igualmente em todas as etapas de construção do artigo: concepção e delineamento do estudo; coleta, análise e interpretação dos dados; redação e revisão crítica do manuscrito; e aprovação da versão final.

Agradecimentos. Agradecemos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Código Financeiro 001, por financiar parte do estudo.

Financiamento. Os autores declaram que tiveram financiamento parcial da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Código Financeiro 001.

Disponibilidade de dados e materiais. Os autores não declararam a disponibilidade dos dados e materiais da pesquisa.

Conflitos de interesse. Os autores não declararam conflitos de interesse.

Declaração de uso de IA. Os autores declaram que ferramentas de IA generativa foram utilizadas exclusivamente como apoio à revisão linguística e à organização textual do manuscrito, sem geração de dados, resultados ou conclusões. Os modelos preditivos descritos correspondem a algoritmos clássicos de aprendizado de máquina (SVC-RBF) e visão computacional (YOLO26s), implementados e validados pelos autores, e constituem o objeto científico do estudo. Todo o conteúdo, a interpretação dos resultados e as conclusões são de inteira responsabilidade dos autores.

REFERÊNCIAS

1. Veasey SC, Rosen IM. Obstructive Sleep Apnea in Adults. N Engl J Med. 2019;380(15):1442-9. doi: https://doi.org/10.1056/NEJMcp1816152
2. Duarte RLM, Togeiro SMGP, Palombini LO, Rizzatti FPG, Fagondes SC, Magalhães-da-Silveira FJ, et al. Brazilian Thoracic Association Consensus on Sleep-disordered Breathing. J Bras Pneumol. 2022;48(4):e20220106. doi: https://doi.org/10.36416/1806-3756/e20220106
3. Chattu VK, Manzar MD, Kumary S, Burman D, Spence DW, Pandi-Perumal SR. The Global Problem of Insufficient Sleep and Its Serious Public Health Implications. Healthcare (Basel). 2018;7(1):1. doi: https://doi.org/10.3390/healthcare7010001
4. Tietjens JR, Claman D, Kezirian EJ, De Marco T, Mirzayan A, Sadroonri B, et al. Obstructive Sleep Apnea in Cardiovascular Disease: A Review of the Literature and Proposed Multidisciplinary Clinical Management Strategy. J Am Heart Assoc. 2019;8(1):e010440. doi: https://doi.org/10.1161/JAHA.118.010440
5. Benjafield AV, Ayas NT, Eastwood PR, Heinzer R, Ip MSM, Morrell MJ, et al. Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis. Lancet Respir Med. 2019;7(8):687-98. doi: https://doi.org/10.1016/S2213-2600(19)30198-5
6. Kapur VK, Auckley DH, Chowdhuri S, Kuhlmann DC, Mehra R, Ramar K, et al. Clinical Practice Guideline for Diagnostic Testing for Adult Obstructive Sleep Apnea: An American Academy of Sleep Medicine Clinical Practice Guideline. J Clin Sleep Med. 2017;13(3):479-504. doi: https://doi.org/10.5664/jcsm.6506
7. Lloyd RM, Crawford T, Donald R, Gray DD, Healy WJ, Junna MR, et al. Quality measure for screening for adult obstructive sleep apnea by primary care providers: 2024 update after measure maintenance. J Clin Sleep Med. 2024;20(11):1819-22. doi: https://doi.org/10.5664/jcsm.11294
8. Chung F, Abdullah HR, Liao P. STOP-Bang Questionnaire: A Practical Approach to Screen for Obstructive Sleep Apnea. Chest. 2016;149(3):631-8. doi: https://doi.org/10.1378/chest.15-0903

9. Bazoukis G, Bollepalli SC, Chung CT, Li X, Tse G, Bartley BL, et al. Application of artificial intelligence in the diagnosis of sleep apnea. J Clin Sleep Med. 2023;19(7):1337-63. doi: https://doi.org/10.5664/jcsm.10532
10. Chiu HY, Chen PY, Chuang LP, Chen NH, Tu YK, Hsieh YJ, et al. Diagnostic accuracy of the Berlin questionnaire, STOP-BANG, STOP, and Epworth sleepiness scale in detecting obstructive sleep apnea: A bivariate meta-analysis. Sleep Med Rev. 2017;36:57-70. doi: https://doi.org/10.1016/j.smrv.2016.10.004
11. Davies RJ, Ali NJ, Stradling JR. Neck circumference and other clinical features in the diagnosis of the obstructive sleep apnoea syndrome. Thorax. 1992;47(2):101-5. doi: https://doi.org/10.1136/thx.47.2.101
12. Hutz MJ, Plata DM, LoSavio P, Herdegen J, Zhang Y, Mokhlesi B. The impact of neck circumference on hypoglossal nerve stimulator therapy outcomes. J Clin Sleep Med. 2024;20(11):1755-61. doi: https://doi.org/10.5664/jcsm.11260
13. Schwab RJ, Pasirstein M, Pierson R, Mackley A, Hachadoorian R, Arens R, et al. Identification of upper airway anatomic risk factors for obstructive sleep apnea with volumetric magnetic resonance imaging. Am J Respir Crit Care Med. 2003;168(5):522-30. doi: https://doi.org/10.1164/rccm.200208-866OC
14. Baker E, Chanamolu M, Nieri C, White SF, Brandt J, Gillespie MB. The Effect of Tongue Volume and Adipose Content on Obstructive Sleep Apnea: Meta-analysis & Systematic Review. OTO Open. 2025;9(2):e70067. doi: https://doi.org/10.1002/oto2.70067
15. Guimarães FN, Canadas J, Cunha MG, Durão V, Rosa P, Rabahi M, et al. Scalloped tongue: an additional, accessible and useful tool to detect severe obstructive sleep apnea? Front Sleep. 2025;4:1652532. doi: https://doi.org/10.3389/frsle.2025.1652532
16. Mencar C, Gallo C, Mantero M, Tarsia P, Carpagnano GE, Foschino Barbaro MP, et al. Application of machine learning to predict obstructive sleep apnea syndrome severity. Health Informatics J. 2020;26(1):298-317. doi: https://doi.org/10.1177/1460458218824725
17. Giorgi L, Nardelli D, Moffa A, Iafrati F, Di Giovanni S, Olszewska E, et al. Advancements in Obstructive Sleep Apnea Diagnosis and Screening Through Artificial Intelligence: A Systematic Review. Healthcare (Basel). 2025;13(2):181. doi: https://doi.org/10.3390/healthcare13020181
18. Sapkota R, Cheppally RH, Sharda A, Karkee M. YOLO26: key architectural enhancements and performance benchmarking for real-time object detection. arXiv [Preprint]. 2025. arXiv:2509.25164.
19. American Academy of Sleep Medicine. American Academy of Sleep Medicine [Internet]. Darien (IL): AASM; 2025 [cited 2025 Sep 25]. Available from: https://aasm.org/
20. Da Fonseca AU, Felix JP, Da Silva Vieira G, Fernandes D, Soares F. Automatic Tuberculosis Detection using Binary Pattern of Phase Congruency. In: 2022 International Conference on Computational Science and Computational Intelligence (CSCI). [s.l.]: IEEE; 2022. p. 1646-51. doi: https://doi.org/10.1109/csci58124.2022.00294
21. Tsuiki S, Nagaoka T, Fukuda T, Sakamoto Y, Almeida FR, Nakayama H, et al. Machine learning for image-based detection of patients with obstructive sleep apnea: an exploratory study. Sleep Breath. 2021;25(4):2297-305. doi: https://doi.org/10.1007/s11325-021-02301-7
22. Kunisaki KM, Greer N, Khalil W, Koffel E, Koeller E, MacDonald R, et al. Provider Types and Outcomes in Obstructive Sleep Apnea Case Finding and Treatment: A Systematic Review. Ann Intern Med. 2018;168(3):195-202. doi: https://doi.org/10.7326/M17-2511
23. Huang WC, Lee PL, Liu YT, Chiang AA, Lai F. Support vector machine prediction of obstructive sleep apnea in a large-scale Chinese clinical sample. Sleep. 2020;43(7):zsz295. doi: https://doi.org/10.1093/sleep/zsz295
24. Goldstein CA, Berry RB, Kent DT, Kristo DA, Seixas AA, Redline S, et al. Artificial intelligence in sleep medicine: an American Academy of Sleep Medicine position statement. J Clin Sleep Med. 2020;16(4):605-7. doi: https://doi.org/10.5664/jcsm.8288
25. Wainer J, Cawley G. Nested cross-validation when selecting classifiers is overzealous for most practical applications. Expert Systems with Applications. 2021;182:115222. doi: https://doi.org/10.1016/j.eswa.2021.115222

NOTAS EDITORIAIS

O Journal of Health Informatics é uma publicação científica mantida pela Sociedade Brasileira de Informática em Saúde, associação científica sem fins lucrativos dedicada ao desenvolvimento da Informática em Saúde e da Saúde Digital no Brasil. O processo editorial é conduzido com independência científica, sem interferência institucional, comercial ou de patrocinadores, e segue as boas práticas internacionais de ética em publicação científica.

Este artigo foi avaliado por pares, em sistema duplo-cego, por especialistas da área. A equipe editorial usou ferramentas de inteligência artificial, com supervisão humana, na triagem inicial, na conferência do texto, das referências e dos números e, quando necessário, no redesenho de figuras, quadros e tabelas, aprovado pelos autores na prova. As ferramentas e as salvaguardas estão na Política de uso de inteligência artificial da revista. As decisões editoriais foram tomadas por pessoas.

SBIS Sociedade Brasileira de Informática em Saúde