



CESPU
INSTITUTO UNIVERSITÁRIO
DE CIÊNCIAS DA SAÚDE

Aplicação da inteligência artificial no diagnóstico precoce do cancro oral

Syma Belilita

Dissertação conducente ao Grau de Mestre em Medicina Dentária (Ciclo Integrado)

—

Gandra, junho de 2023

Syma Belilita

**Dissertação conducente ao Grau de Mestre em Medicina Dentária
(Ciclo Integrado)**

Aplicação da inteligência artificial no diagnóstico precoce do cancro oral

Trabalho realizado sob a Orientação de
Professor Doutor José Manuel Barbas do Amaral

DECLARAÇÃO DE INTEGRIDADE

Eu, acima identificado, declaro ter atuado com absoluta integridade na elaboração deste trabalho, confirmo que em todo o trabalho conducente à sua elaboração não recorri a qualquer forma de falsificação de resultados ou à prática de plágio (ato pelo qual um indivíduo, mesmo por omissão, assume a autoria do trabalho intelectual pertencente a outrem, na sua totalidade ou em partes dele). Mais declaro que todas as frases que retirei de trabalhos anteriores pertencentes a outros autores foram referenciadas ou redigidas com novas palavras, tendo neste caso colocado a citação da fonte bibliográfica.

Agradecimentos

Gostaria de agradecer aos meus pais e minha irmã pelo seu apoio incondicional, amor e encorajamento constantes. A sua presença e os seus sacrifícios têm sido uma fonte de inspiração para mim.

Gostaria também de agradecer à pessoa que partilha a minha vida, pelo seu suporte, compreensão e presença, que tem sido um pilar ao meu lado durante todos estes anos.

Um agradecimento especial a minha parceira e companheira de casa, Özgül, obrigada por todas as noites de estudo, risos memoráveis e momentos de amizade. Obrigada por todas as lembranças preciosas que partilhámos juntos nos últimos 5 anos. Estou grata por ter podido contar contigo em cada passo.

E, finalmente, aos meus amigos em Portugal, Inara, Sybel, Alice, Roméïssa, Adel, e em França, obrigada por me terem apoiado e encorajado ao longo destes anos. Os vossos sorrisos, os vossos conselhos e a vossa amizade sincera fizeram desta uma aventura inesquecível.

Gostaria também de expressar a minha mais profunda gratidão ao meu professor de tese, que me orientou com sabedoria, perícia e paciência. Os seus conselhos esclarecidos e a sua paixão pela investigação foram uma fonte constante de inspiração.

A todos vós, estou infinitamente grata por terem feito parte do meu percurso.

Resumo

INTRODUÇÃO: O diagnóstico e a gestão das lesões orais malignas são desafios importantes na odontologia, especialmente devido à falta de triagem e vigilância eficazes, levando a diagnósticos tardios e baixas taxas de sobrevivência do cancro oral (CO). Esta revisão explora o uso da inteligência artificial (IA) para a deteção precoce do CO, abordando suas vantagens e impacto na gestão da doença.

OBJETIVO: Explorar o potencial e papel da IA no diagnóstico precoce do CO e analisar suas vantagens, limites e recomendações.

MATERIAIS E MÉTODOS: Foi conduzida uma pesquisa bibliográfica nas bases de dados PubMed, Pubmed US Central, B-ON, e eClinical Medecine. Utilizou-se a funcionalidade de pesquisa avançada, combinando as palavras-chave "*oral cancer*", "*algorithm*", "*machine learning*" e "*artificial intelligence*", "*cancer screening*".

RESULTADOS: Após análise de 226 artigos, 43 deles atenderam aos critérios de inclusão com base nos títulos e resumos. Após uma leitura completa, 20 artigos foram considerados irrelevantes para o estudo. Associado a seleção de 16 artigos para apoio teórico, um total de 38 artigos foram incluídos nesta revisão sistemática integrativa.

DISCUSSÃO: Os estudos concordam que a IA é uma abordagem promissora para a deteção precoce do CO. No entanto, subsistem limitações, como a necessidade de dados de qualidade, a necessidade de validação externa e os desafios éticos associados à utilização da IA. São necessários mais esforços para ultrapassar estas limitações e maximizar o potencial da IA.

CONCLUSÃO: Com uma abordagem responsável e uma colaboração multidisciplinar, a IA pode desempenhar um papel crucial na melhoria dos resultados clínicos e da sobrevivência dos doentes com CO, auxiliando os médicos-dentistas.

PALAVRAS-CHAVES: "*oral cancer*", "*algorithm*", "*machine learning*", "*artificial intelligence*" e "*cancer screening*"

Abstract

INTRODUCTION: The diagnosis and management of oral malignant lesions are major challenges in dentistry, especially due to lack of effective screening and surveillance, leading to late diagnosis and low survival rates of oral cancer (OC). This review explores the use of artificial intelligence (AI) for early detection of OC, addressing its advantages and impact on disease management.

OBJECTIVE: To explore the potential and role of AI in the early diagnosis of OC and to analyze its advantages, limits and recommendations.

MATERIALS AND METHODS: A literature search was conducted in PubMed, Pubmed US Central, B-ON, and eClinical Medecine databases. Advanced search functionality was used, combining the keywords "oral cancer", "algorithm", "machine learning" and "artificial intelligence", "cancer screening".

RESULTS: After analysis of 226 articles, 43 of them met the inclusion criteria based on titles and abstracts. After a thorough reading, 20 articles were considered irrelevant to the study. Associated with the selection of 16 articles for theoretical support, a total of 38 articles were included in this integrative systematic review.

DISCUSSION: The studies agreed that AI is a promising approach for the early detection of OC. However, limitations remain, such as the need of high-quality data, the need of external validation and the ethical challenges associated with the use of AI. Further efforts are needed to overcome these limitations and maximize the potential of AI.

CONCLUSION: With a responsible approach and multidisciplinary collaboration, AI can play a crucial role in supporting the dentists and improving clinical outcomes and survival of patients with OC.

KEY WORDS: "oral cancer", "algorithm", "machine learning" , "artificial intelligence" and "cancer screening"

Índice geral

1. INTRODUÇÃO.....	1
2. OBJETIVOS.....	3
2.1. OBJETIVO GERAL.....	3
2.2. OBJETIVOS ESPECÍFICOS.....	3
3. MATERIAIS E MÉTODOS	3
3.1. METODOLOGIA DA CONSTRUÇÃO DA INTELIGÊNCIA ARTIFICIAL.....	3
3.2. PROTOCOLO DESENVOLVIDO	6
3.3. CRITÉRIOS DE ELEGIBILIDADE POR MÉTODO PICO (PERGUNTA QUALITATIVA).....	7
3.4. CRITÉRIOS DE INCLUSÃO E EXCLUSÃO.....	8
3.5. ESTRATÉGIA DE PESQUISA	8
4. RESULTADOS.....	11
4.1. EXTRAÇÃO DE DADOS	11
4.2. CARACTERÍSTICAS DOS ESTUDOS	12
4.3. CARACTERÍSTICAS DOS MODELOS DE IA INVESTIGADOS NOS ESTUDOS.....	12
4.4. CLASSIFICAÇÃO DA QUALIDADE DAS REVISTAS ACADÉMICAS PELO <i>SCIENTIFIC JOURNAL RANKING</i> (SJR).....	14
5. DISCUSSÃO.....	37
5.1. PRINCIPAIS IAs UTILIZADAS E SUAS VANTAGENS	37
5.2. MÉTODOS DE AQUISIÇÃO DE DADOS	41
6.2.A. <i>Aplicações clínicas e hospitalares.....</i>	<i>41</i>
6.2.B. <i>Aplicáveis a práticas ou áreas com poucos recursos.....</i>	<i>45</i>
5.3. DESAFIO DA QUALIDADE DOS DADOS E DA UTILIZAÇÃO DE VARIÁVEIS.....	48
5.4. LIMITES, ÉTICA E CONTEXTO ATUAL.....	51
6. CONCLUSÕES.....	57
7. REFERENCIAS BIBLIOGRÁFICAS	58

Índice de figuras

<i>Figura 1: Explicação do funcionamento de machine e deep learning de acordo com Ilhan et al. (2020) (13).....</i>	<i>4</i>
<i>Figura 2: Fluxograma da pesquisa bibliográfica.....</i>	<i>11</i>
<i>Figura 3: Representação da utilização dos diferentes tipos de IA de acordo com os 23 artigos selecionados (3,7,10,15–34).....</i>	<i>13</i>
<i>Figura 4: Comparação do desempenho humano e da IA de acordo com Warin et al., Heo et al., Yuan et al., He et al., Fu et al. e Wang et al. (3,16,17,21,26,27).....</i>	<i>13</i>

Índice de tabela

<i>Tabela 1: Métricas aplicadas para avaliar o desempenho dos modelos de IA dos artigos no conjunto de teste de acordo com Jubair et al. (15)</i>	6
<i>Tabela 2: Pergunta de interesse com base PICO</i>	7
<i>Tabela 3: Critérios de elegibilidade (inclusão e exclusão)</i>	8
<i>Tabela 4: Resultados da estratégia de busca</i>	9
<i>Tabela 5: Classificação da qualidade das revistas académicas pelo SJR</i>	14
<i>Tabela 6: Modelos de inteligência artificial com base de dados os fatores de risco e características clínico-patológicas</i>	15
<i>Tabela 7: Modelos de inteligência artificial com base de dados de fotografias orais</i>	17
<i>Tabela 8: Modelos de inteligência artificial com base de dados adquirida com smartphone</i>	22
<i>Tabela 9: Modelos de inteligência artificial com base de dados amostras salivares</i>	25
<i>Tabela 10: Modelos de inteligência artificial com base de dados de técnicas de imagiologia</i>	27
<i>Tabela 11: Modelos de inteligência artificial com base de dados imagens de tomografia de coerência ótica (TCO)</i>	32
<i>Tabela 12: Modelos de inteligência artificial com base de dados de análise de biomarcadores</i>	34

Lista de abreviaturas, siglas e acrónimos

CO – Cancro oral

DPMO - Distúrbios potencialmente malignos orais

IA – Inteligência artificial

ML- *Machine learning*

SE – Sensibilidade

EP – Especificidade

ASC – Área sob a curva

CECO - Carcinoma espinocelular oral

TGS - Tumor das glândulas salivares

TGP - Tumores da glândula parotídea

RMN – Ressonância magnética nuclear

TCO – Tomografia de coerência ótica

RNA – Ácido ribonucleico

PCR - *Polymerase Chain Reaction* (reação de polimerização em cadeia)

KNN - K-vizinhos mais próximos

FA - Floresta Aleatória

MVS - Máquina Vetorial de Suporte

ADL - Análise Discriminatória Linear

RL - Regressão Logística

AD - Árvore de decisão

RCF - Rede de Crença Profunda

FBR - Função de base radial

RNC – Rede neural convolucional

SJR - *Scientific Journal Ranking*

RNA - Rede neural artificial

VPP - Valor preditivo positivo

DL – *Deep learning*

SSD - *Single Shot MultiBox Detector*

FWH - *Frontline-health-worker*

ILB - Imagem de Luz Branca

IAF - Imagem de Autofluorescência

NB - Naive Bayes

ECM - Espaço de Calibração Multidimensional

MVL - Máquina Vetorial de Suporte com núcleos linear

MVP - Máquina Vetorial de Suporte com núcleos: linear polinomial

MVR - Máquina Vetorial de Suporte com núcleos: linear radial

ELISA - *Enzyme-linked immuno-sorbent assay*

LASSO - *Least Absolute Shrinkage and Selection Operator*

DWI - *Diffusion-weighted Imaging*

VM - Votação por maioria

ECL - Endomicroscopia confocal a laser

AP - Adenoma pleomórfico

ACB - Adenoma de células basais,

CM - Carcinoma mucoepidermóide,

CAC - Carcinoma adenóide cístico

PCA - *Principal Component Analysis*

ARPM - Aprendizagem residual profunda multinível

OMS – Organização mundial da saúde

HRNet - *High-Resolution Network*

VGGNet - *Visual Geometry Group Network*

ResNet - *Residual Network*

1. Introdução

“O câncer oral é uma doença caracterizada pelo crescimento anormal e descontrolado de células malignas nos tecidos da boca, incluindo lábios, língua, gengivas, palato, bochechas e garganta. Esse tipo de câncer pode se espalhar para outras áreas do corpo e pode ser fatal se não for tratado precocemente. ”

CHATGTP

Um dos maiores desafios que os dentistas enfrentam atualmente é a identificação e gestão das lesões orais devido à falta de experiência, recursos ou atrasos na procura de cuidados médicos. De facto, o cancro oral (CO) é caracterizado por uma das menores taxas de sobrevivência mundial. De acordo com o GLOBOCAN, foram registados 382 óbitos em 1103 casos diagnosticados em Portugal que corresponde a uma taxa de sobrevivência de 65 %.(1) A principal causa de maus prognósticos de CO é uma triagem ausente e vigilância ineficazes ao nível local de atendimento que é feita pelo médico dentista levando a atrasos de diagnóstico.(2) As fases iniciais podem ser difíceis de identificar uma vez que são frequentemente assintomáticas e podem assemelhar-se a lesões benignas.(3) Doentes com menor intervalo desde a deteção inicial de uma lesão suspeita até o diagnóstico histológico são mais suscetíveis de ter um melhor prognóstico do que aqueles com diagnóstico tardio.(2) A origem desse atraso de diagnóstico é discutida. A disseminação metastática pode já ter ocorrido antes que o tumor atinja um tamanho mensurável e significativo para o doente.(4) A demora no diagnóstico pode ser causada tanto pelo atraso dos doentes quanto pelo dos dentistas, o que torna o termo "diagnóstico precoce" confuso.(5)

Para uma melhor prevenção, a Organização Mundial de Saúde divulgou um relatório que enumera os distúrbios potencialmente malignos orais (DPMOs). Estes DPMOs estão relacionados a um risco aumentado de ocorrência do cancro, o que pode, portanto, aumentar o número de patologias identificáveis para o diagnóstico da doença. (6) Com efeito, quando as taxas de sobrevivência a 5 anos são de apenas 45% e 32%,

respetivamente, dois terços dos DPMOS são identificados na fase III ou na fase IV. A taxa de sobrevivência sobe para 84% para o terço restante que são descobertos nas fases iniciais. (7)

Para lidar com estes assuntos, o mundo da saúde está passando por mudanças sem precedentes graças aos avanços tecnológicos recentes. No centro das notícias, desenvolve-se uma nova tecnologia, divulgada ao grande público na sua forma mais conhecida ChatGTP, a inteligência artificial (IA).

IA se refere à capacidade de as máquinas simularem a inteligência e o comportamento humano, por meio de diversas tecnologias emergentes, cada vez mais presentes no nosso dia a dia. Uma dessas tecnologias é o *machine learning* (ML), que permite que os computadores desenvolvam capacidades de resolução de problemas e aprendizagem sem a necessidade de serem programados explicitamente. (8) Essa tecnologia traz benefícios significativos para profissionais de saúde, sistemas de saúde e doentes, permitindo análise e interpretação de grandes volumes de dados em questão de segundos, auxiliando no processo de tomada de decisão.(9)

Um exemplo concreto de aplicação da IA pode ser a deteção precoce do câncer oral, por meio da análise de imagens e dados médicos. Atualmente, o diagnóstico padrão é baseado na deteção visual de lesões suspeitas, seguido de biópsia invasiva. (10) No entanto, a IA pode ser aplicada para analisar imagens orais e identificar áreas anormais que requerem atenção, permitindo o desenvolvimento de uma solução não invasiva e de fácil utilização para os médicos.

Esta revisão pretende reunir evidências científicas claras que comprovam que a medicina digital, por meio da IA, tem potencial para mudar as diretrizes na gestão do CO e orientar a decisão terapêutica, proporcionando maior eficiência e precisão no diagnóstico e tratamento.

2. Objetivos

2.1. Objetivo geral

O objetivo desta revisão sistemática integrativa consiste em analisar o estado atual da investigação sobre a aplicação da inteligência artificial (IA) na deteção precoce do cancro oral.

2.2. Objetivos específicos

A revisão centra-se na identificação das diferentes abordagens e métodos utilizados, na avaliação da validade dos estudos existentes, na síntese dos resultados e na formulação de recomendações práticas. Os objetivos são obter uma melhor compreensão da eficácia da IA neste domínio, identificar lacunas na investigação e propor orientações para a utilização ética e prática da IA.

3. Materiais e métodos

3.1. Metodologia da construção da inteligência artificial

De forma a melhorar a compreensão e a visualização do conceito de inteligência artificial, é necessária uma fundamentação teórica dos termos.

Machine learning, um ramo da inteligência artificial que se centra na utilização de algoritmos para resolver desafios como a categorização de dados, é uma área de interesse crescente para os investigadores que procuram transformar grandes quantidades de dados em conhecimentos que possam ser úteis para a tomada de decisões clínicas. (11)

As diferentes abordagens algorítmicas de inteligência artificial mais comumente utilizadas no domínio da saúde são as máquinas de vetores de apoio (MVS), as árvores de decisão (AD) e as redes neuronais. (12)

Um ramo específico da aprendizagem automática desenvolveu-se consideravelmente nos últimos anos, o *Deep Learning*, que utiliza redes neuronais profundas, incluindo as redes neurais convolucionais (RNC), com várias camadas ocultas para aprender representações hierárquicas dos dados. (13)

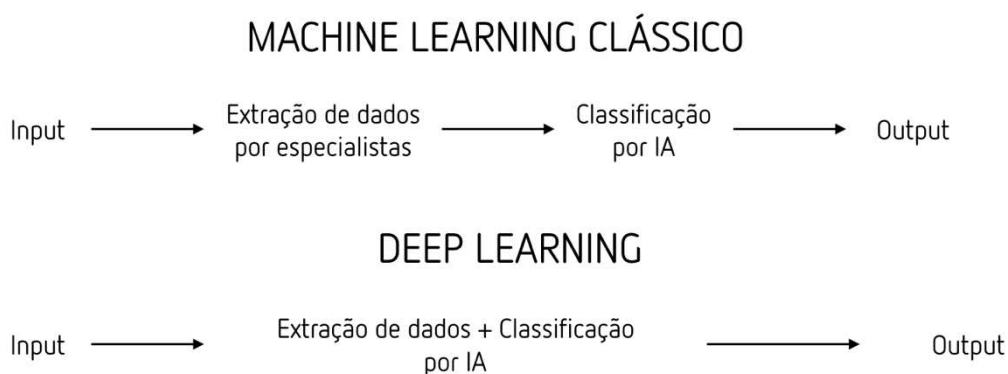


Figura 1: *Explicação do funcionamento de machine e deep learning de acordo com Ilhan et al. (2020) (13)*

As redes neuronais artificiais, que são constituídas por camadas de neurónios interligados, são modeladas de acordo com o funcionamento do cérebro humano. Cada neurónio de uma rede neuronal aplica matemática aos dados de entrada antes de enviar o resultado para a camada seguinte. Ao alterar os pesos das ligações entre os neurónios, as redes neuronais podem aprender com os dados de treino e prever os resultados de novos dados.

A característica distintiva da aprendizagem profunda consiste na extensão das redes neuronais, ou seja, no número de camadas intermédias entre as camadas de entrada e de saída. Ao utilizar redes neuronais profundas, torna-se possível aprender representações cada vez mais abstratas e complexas dos dados à medida que a informação se propaga através das diferentes camadas da rede.

Para o desenvolvimento do modelo, é necessária uma fase de formação. Um conjunto de dados específicos desta etapa é introduzido no modelo de IA durante a fase de treino. Como o modelo aprende a associar as entradas aos resultados associados, os dados de treino consistem normalmente em pares de entradas e saídas esperadas. Nesta fase, os pesos e os parâmetros do modelo são ajustados para o ensinar a proceder a extrações a partir de determinados exemplos. (14)

Esta abordagem, em que um modelo de IA é treinado utilizando dados marcados, ou seja, dados que foram previamente anotados com etiquetas ou resultados esperados, frequentemente por especialistas, até atingir um desempenho aceitável é designada por inteligência artificial supervisionada. (14)

Previamente ao treino, existem três fases fundamentais na operação dos dados: pré-processamento, segmentação e pós-processamento dos dados.

O pré-processamento prepara os dados antes da aprendizagem, incluindo a remoção do ruído das imagens e a gestão dos valores atípicos, a segmentação identifica e divide as partes significativas dos dados e o pós-processamento refina e melhora os resultados da IA. Estas etapas são cruciais para obter um desempenho ótimo e uma compreensão precisa dos dados ao longo das várias fases do processamento da IA. São utilizadas diferentes abordagens para tratar os dados, algumas das quais serão discutidas no nosso estudo. (13)

Após a fase de formação e processamento, é necessário avaliar as capacidades do algoritmo de aprendizagem.

Esta etapa, conhecida como a fase de teste, envolve a utilização de um conjunto de dados de avaliação separado, que não foi utilizado durante a formação, para testar o desempenho do modelo em novos dados. Durante o teste, o modelo faz previsões sobre os exemplos de dados de teste e compara essas previsões com os resultados esperados.

Foram aplicadas várias medidas de desempenho nos estudos, e consideraremos a precisão, a sensibilidade (SE), a especificidade (EP) e a área sob a curva (ASC). (15)

Sensibilidade	Mede a capacidade de um modelo para identificar os verdadeiros positivos entre todos os sujeitos positivos = Proporção de verdadeiros positivos corretamente identificados pelo modelo.
Especificidade	Mede a capacidade de um modelo para identificar os verdadeiros negativos entre todos os sujeitos negativos = Proporção de verdadeiros negativos corretamente identificados pelo modelo.
Precisão	Mede a capacidade de prever corretamente exemplos positivos e negativos = Rácio entre o número total de previsões corretas e o número total de exemplos no conjunto de dados.
Área sob a curva	Mede a capacidade de um modelo para classificar corretamente exemplos positivos e negativos, independentemente do limiar de classificação utilizado. Quanto mais próximo de 1, melhor será a capacidade do modelo para distinguir as classes .

Tabela 1: Métricas aplicadas para avaliar o desempenho dos modelos de IA dos artigos no conjunto de teste de acordo com Jubair et al. (15)

Quanto mais próximos de 1 (ou de 100%) estiverem os valores da sensibilidade, especificidade, precisão e área sob a curva, melhor é o desempenho do algoritmo de aprendizagem. (15)

3.2. Protocolo desenvolvido

Para realizar esta revisão sistemática integrativa, foi estabelecido um protocolo de acordo com as diretrizes da declaração PRISMA 2020, que é um guia de referência amplamente reconhecido para revisões sistemáticas. Esse protocolo orientou a seleção dos estudos, a extração de dados e a análise dos resultados, garantindo assim a transparência e a rigor metodológica do estudo.

3.3. Critérios de elegibilidade por método PICO (pergunta qualitativa)

Considerando o objetivo desta revisão sistemática, a seguinte pergunta de pesquisa foi formulada com base na estratégia PICO:

“Na deteção precoce do cancro oral, a aplicação da inteligência artificial traz benefícios aos protocolos atuais de diagnóstico?”

As perguntas de interesse com base na população estudada (P), o interesse da técnica estudada (I), a comparação (C), e *outcome* (O) são expostas na tabela seguinte (Tabela 1).

População (P)	Indivíduos com risco de cancro oral ou diagnosticados com cancro oral
Interesse (I)	Utilização de diferentes modelos de inteligência artificial para a deteção precoce do cancro oral
Comparação (C)	Comparação do desempenho de diferentes modelos de inteligência artificial entre si e com os atuais protocolos de deteção do cancro
<i>Outcome</i> (O)	A mais-valia que IA proporcione na deteção precoce do cancro oral

Tabela 2: *Pergunta de interesse com base PICO*

3.4. Critérios de inclusão e exclusão

Critérios de inclusão	Critérios de exclusão
<ul style="list-style-type: none"> ⇒ Língua inglês, português e francês ⇒ Data de publicação não limitada ⇒ Doentes adultos ⇒ Estudos que fornecem dados sobre o diagnóstico precoce ⇒ Estudos de diagnóstico, comparativos, prospetivos e retrospectivos ⇒ Estudos que avaliam modelos de aprendizagem automático 	<ul style="list-style-type: none"> ⇒ Artigos não acessíveis ou incompletos ⇒ Dissertações, revistas sistemáticas ⇒ Artigos cujos títulos não tratam do cancro oral ⇒ Artigos cuja leitura na íntegra não fornecem informações necessárias sobre o assunto ⇒ Artigos que tratam da IA no seu uso nas diferentes áreas ⇒ Artigos que trataram da taxa de sobrevivência do cancro oral

Tabela 3: *Critérios de elegibilidade (inclusão e exclusão)*

3.5. Estratégia de pesquisa

A seleção de artigos para esta revisão sistemática integrativa seguiu uma abordagem metodológica baseada nos critérios da declaração PRISMA. Foi conduzida uma extensa pesquisa bibliográfica em 4 bases de dados diferentes, incluindo PubMed, Pubmed US Central, B-ON, e eClinical Medecine. Utilizou-se a funcionalidade de pesquisa avançada, combinando as palavras-chave relevantes, tais como "oral cancer", "algorithm", "machine learning", "artificial intelligence" e "cancer screening". Essa abordagem meticulosa garantiu a abrangência e a qualidade dos estudos considerados para análise nesta pesquisa. A estratégia de busca está ilustrada na tabela seguinte (Tabela 3).

Base de dados	Estratégia de busca	Artigos identificados	Artigos selecionados
PubMed	(« oral cancer »[MeSH Terms]) AND (« algorithm»[MeSH Terms]) AND (« machine learning» [MeSH Terms])	118	16
PubMed	(« oral cancer »[MeSH Terms]) AND (« cancer screening »[MeSH Terms]) AND (« ai artificial intelligence »[MeSH Terms])	21	1
PubMed US Central	(« oral cancer »[MeSH Terms]) AND (« machine learning »[MeSH Terms])	44	2
B-ON	SU (oral cancer or carcinoma in oral cavity or malignancy in oral cavity) AND SU (artificial intelligence) AND SU (deep learning)	32	3
eClinical Medecine	(oral cancer) and (artificial intelligence)	11	1

Tabela 4: *Resultados da estratégia de busca*

Os resultados desta pesquisa, que incluiu um total de 226 artigos, foram importados para o software de gestão de referências Mendeley.

A primeira fase de seleção envolveu a eliminação de artigos duplicados detetados através da funcionalidade do Mendeley, resultando na remoção de 55 artigos. Em seguida, os títulos e resumos dos restantes artigos foram avaliados quanto à sua relevância preliminar para os objetivos desta revisão, permitindo uma seleção inicial. Após a análise dos títulos, foram retidos 73 artigos e, após a análise dos resumos, foram selecionados 43 artigos para estudo posterior.

Os 43 artigos pré-selecionados foram depois submetidos a uma avaliação mais aprofundada, através da análise do seu conteúdo integral. Foram eliminados 20 artigos por não cumprirem os critérios de elegibilidade enumerados na Tabela 2.

Após esta avaliação, foram escolhidos 23 artigos explorados nos resultados pela sua relevância e pela sua capacidade de fornecer informações significativas em resposta à questão de investigação da revisão sistemática.

De referir que, de forma a enriquecer a introdução e discussão da revisão, foi efetuada uma pesquisa manual para identificar outros estudos relevantes que pudessem complementar a revisão. Após uma avaliação cuidadosa da sua relevância, foram incluídos na seleção 15 artigos adicionais, incluindo uma proposta de lei, um relatório GLOBOCAN, 2 artigos editoriais, 9 revisões da literatura, uma revisão sistemática e um livro.

Os esquemas realizados para as figuras 2,3 e 4 são uma realização pessoal.

4. Resultados

4.1. Extração de dados

Foi elaborado um fluxograma para ilustrar visualmente as diferentes fases do processo de seleção dos artigos. (Figura 1) Este fluxograma confere transparência e clareza metodológica à revisão.

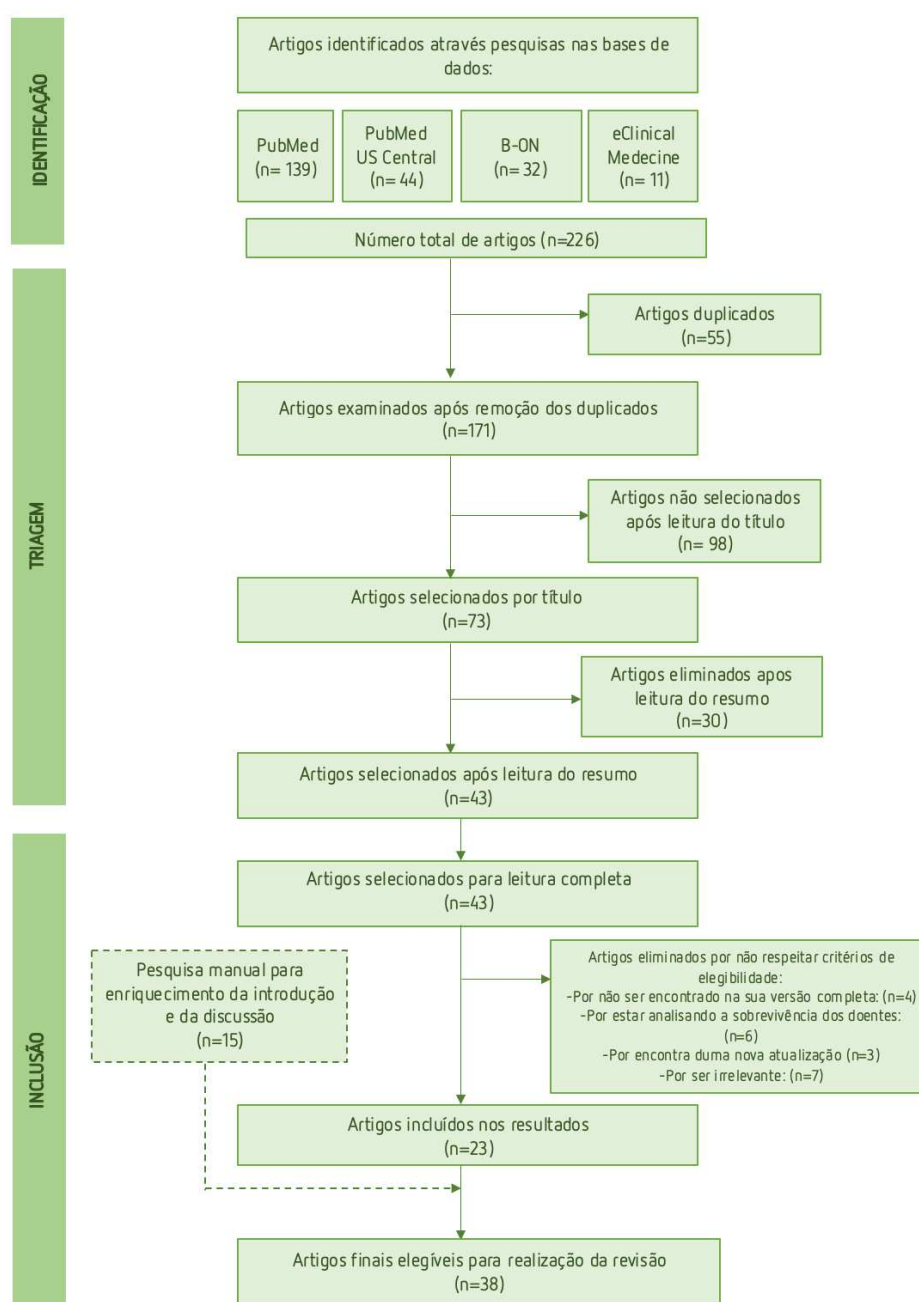


Figura 2: Fluxograma da pesquisa bibliográfica

4.2. Características dos estudos

De acordo com os critérios de inclusão e exclusão, a pesquisa selecionou 23 artigos que contribuíram para o objetivo do estudo. Das 23 publicações, 11 investigações eram estudos de diagnóstico comparativo, um era um estudo estritamente de diagnóstico, 7 eram estudos retrospectivos, 3 eram estudos prospectivos e um era um estudo observacional prospectivo.

Para proceder à investigação, os artigos estudaram vários cancros orais, 7 tratavam de cancros orais em geral, 10 de carcinoma espinocelular oral (CECO), 4 de tumores das glândulas salivares (TGS), 2 mais especificamente de tumores da glândula parótida (TGP), 2 do cancro da língua, 1 do cancro do pavimento bucal e 3 de distúrbios potencialmente malignos orais (DPMOs).

Para a deteção do cancro oral, foram estudadas diferentes amostras, 2 estudos tiveram em conta os dados clínicos dos doentes, 6 utilizaram fotografias orais como base de dados, 3 imagens de smartphones para áreas com poucos recursos, 2 imagens de luz branca e 1 imagens de fluorescência.

Outras investigações exploraram dados de imagiologia, 3 dados de ressonância magnética nuclear (RMN), 2 imagens de tomografia de coerência ótica (TCO), 1 outro artigo examina imagens hiperespectrais, 1 imagens de endoscopia e uma outra publicação, imagens de endomicroscopia confocal a laser.

Alguns artigos também investigaram diferentes tipos de biomarcadores, 2 artigos analisaram amostras salivares, um sequências de miRNA (ácido ribonucleico), um estudo amostras citológicas e um estudo considerou amostras respiratórias.

4.3. Características dos modelos de IA investigados nos estudos

Para ilustrar a dispersão da utilização dos diferentes algoritmos de aprendizagem, foi efetuada uma classificação destes modelos associada à sua frequência de utilização nos artigos encontrados nos resultados.

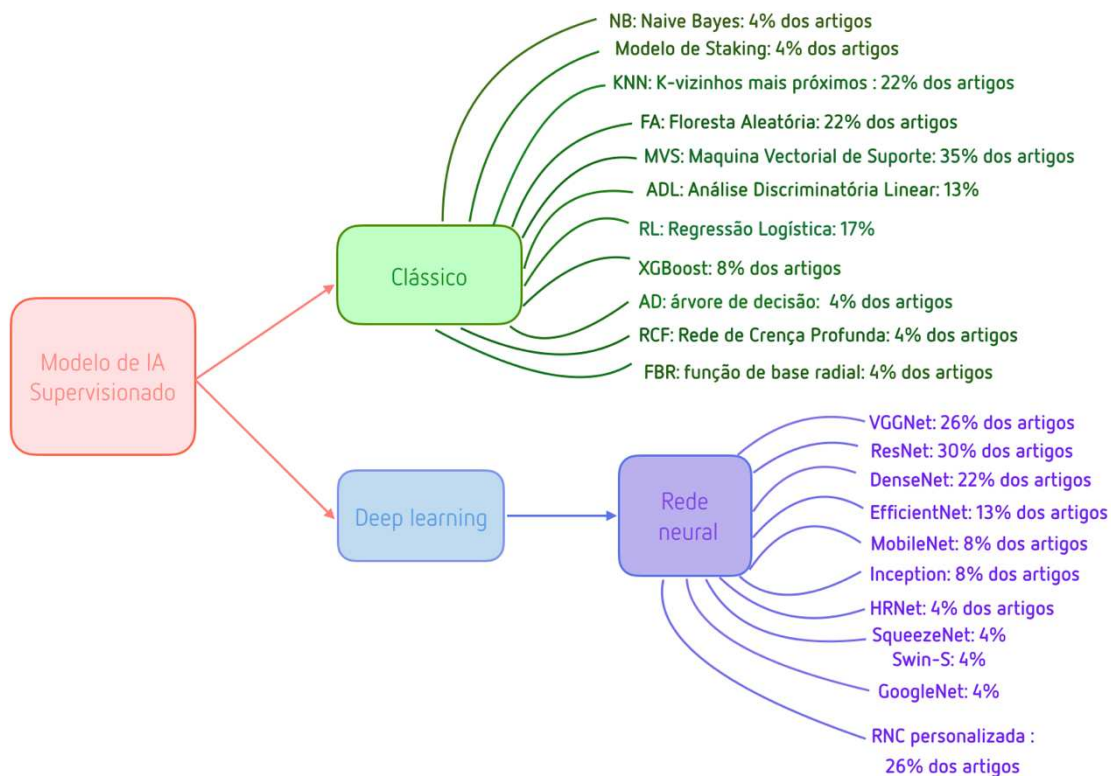


Figura 3: Representação da utilização dos diferentes tipos de IA de acordo com os 23 artigos selecionados (3,7,10,15–34)

Para medir o desempenho real dos modelos de aprendizagem, vários artigos também tiveram em conta o desempenho humano com especialistas e médicos dentistas generalistas.

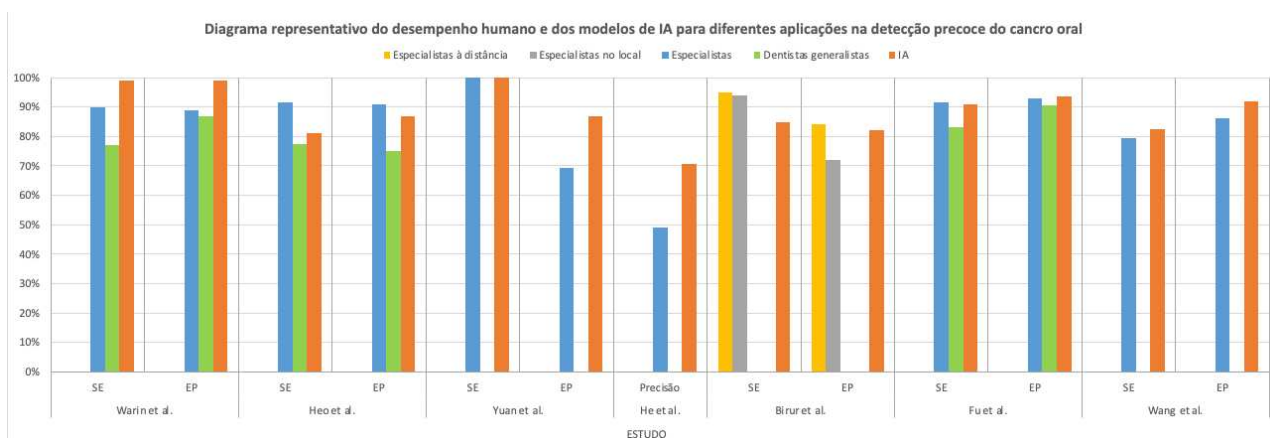


Figura 4: Comparação do desempenho humano e da IA de acordo com Warin et al., Heo et al., Yuan et al., He et al., Fu et al. e Wang et al. (3,16,17,21,26,27)

Dos 6 artigos que comparam o desempenho humano, 4 consideram que a IA supera os especialistas (3,17,26,27), 1 artigo considera que os especialistas e a IA obtêm resultados equivalentes (16) e 1 artigo considera que os especialistas superam o algoritmo de aprendizagem (21). Todos os 6 artigos concordam que a IA supera os profissionais de saúde generalistas.

4.4. Classificação da qualidade das revistas académicas pelo *Scientific Journal Ranking (SJR)*

Estudo	Revista académica	Classificação
Alhazmi A, et al, 2021	Journal of oral pathology & medicine	Q3
Wang X, et al, 2020	Journal of oral pathology & medicine	Q3
Fu Q et al, 2020	EClinical Medicine	Q1
Ferrer Sánchez A et al, 2022	Oral oncology	Q1
Lin H et al, 2021	Journal of biomedical optics	Q2
Warin K et al, 2022	PLoS ONE	Q1
Marzouk R et al, 2022	Computers, Materials and Continua	Q2
Song B et al, 2021	Journal of biomedical optics	Q2
Birur N et al, 2022	Scientific reports	Q1
Jubair F et al, 2021	Oral Diseases	Q1
Braz D et al, 2022	Talanta	Q1
Tseng Y et al, 2022	BMC Oral Health	Q1
Shao S et al, 2020	Journal of X-Ray Science and Technology	Q3
Gunduz E et al, 2022	European Archives of Oto-Rhino-Laryngology	Q2
He Z et al, 2022	European Radiology	Q1
Heo J et al, 2022	Scientific Reports	Q1
Hsieh C et al, 2018	Scientific Reports	Q1
Jeyaraj P et al, 2019	Journal of cancer research and clinical oncology	Q2
McRae M et al, 2020	Cancer cytopathology	Q2
Mentel S et al, 2021	BMC Oral Health	Q1
Aubreville M et al, 2017	Scientific Reports	Q1
Yang Z et al, 2022	Lasers in Medical Science	Q2
Yuan W et al, 2022	Oral Diseases	Q1

Tabela 5: *Classificação da qualidade das revistas académicas pelo SJR*

Dos 19 jornais estudados como fontes, todos foram identificados na base de dados do SJR. Dos nossos 23 artigos, 57% tinham uma classificação Q1 com alta qualidade, 30% uma classificação Q2 com boa qualidade e 13% uma classificação Q3 com qualidade média. Os artigos citados nesta investigação são, portanto, maioritariamente retirados de revistas científicas prestigiadas e influentes, com numerosas citações.

Tabela 6: Modelos de inteligência artificial com base de dados os fatores de risco e características clínico-patológicas

Autor/Ano/Pais	Tipo de estudo	Amostra	Tipo de cancro	Método de classificação	Objetivo	Variáveis	Resultados	Conclusão
Alhazmi A, Alhazmi Y, Makrami A, et al Alhazmi A, Alhazmi Y, Makrami A et al 2021 Arábia saudita	Estudo diagnóstico	73 doentes:22 casos benignos e 51 casos malignos do Hospital do Príncipe Mohammed Bin Nasser entre 2017 e 2020 dos quais 37 mulheres e 36 homens 54 usados para treino e 19 para fase de teste ¹	Carcinoma espino-celular oral (CECO)	Algoritmo de extração de dados: Rede neural artificial (RNA modelo)	Desenvolver um modelo de rede neural artificial que ajude a prever o risco dos indivíduos de desenvolverem cancro oral com base de dados fatores de risco, condição médica sistémica, e características clínico-patológicas.	29 variáveis (idade, queixa principal, condição médica (diabetes, hipertensão), localização, características clínico-patológicas, fatores de risco (tabaco, tabaco mascado, khat)	Na análise de validação cruzada de 10 vezes: <ul style="list-style-type: none"> Sensibilidade (SE) média = 85,71% (intervalo de confiança de 95% [IC], 57,19-98,22) Especificidade (EP) = 60,00% (IC 95%, 14,66-94,73), A precisão do RNA para a previsão do cancro oral foi de 78,95% (95% CI, 54,43-931,95).	Os resultados mostraram que o modelo RNA poderia bem estimar a probabilidade de malignidade e melhorar o valor preditivo positivo (VPP), o que poderia fornecer aos prestadores de cuidados de saúde uma ferramenta simples para ajudar a prever o risco de malignidade, com base o conhecimento dos fatores de risco, condições médicas sistémicas e dados clínicos e patológicos para pessoas com potencial cancro oral

CECO : Carcinoma espinocelular oral, RNA: Rede neural artificial, SE: sensibilidade, EP: especificidade, VPP: valor preditivo positivo

<p>Wang X, Yang J, Wei C et al. 2020 China</p>	<p>Estudo diagnóstico prospetivo</p>	<p>266 doentes do Departamento de Medicina Oral do Hospital de Estomatologia da China Ocidental, de Junho de 2013 a Dezembro de 2016 com DPMOs</p> <p>Fase de teste: 38 participantes da Universidade de Wuhan com DPMOs</p> <p>101 participantes: seguimento de 3 anos</p>	<p>CO, DPMOs</p>	<p>Floresta Aleatória (FA)</p> <p>Modelo B: apenas exame oral não invasivo (fluorescência, azul de toluidina)</p> <p>Modelo P: exame oral não invasivo e características pessoais (informações sobre a lesão, idade, sexo, hábitos)</p> <p>Especialistas</p> <p>Sistema de classificação da fluorescência: verde (0), cinzento (1), preto (2)</p> <p>Classificação coloração com azul de toluidina: sem coloração (0), azul-pálido (1) e azul-escuro (2)</p>	<p>Criar um modelo computacional personalizado para estimar o nível de risco de cancro das DPMOs e investigar a sua possível utilização na internet para o rastreio de DPMOs</p>	<p>Idade, sexo, local da lesão, textura, infiltração, tamanho, queratinização e tipo clínico, local da lesão visualmente melhorado com fluorescência e coloração com azul de toluidina, história de tabagismo, história de consumo de álcool, história de alimentos condimentados</p>	<p>Modelo P otimizado: os testes não invasivos, combinados com 4 preditores pessoais (infiltração da lesão, seu tipo clínico e local, e idade)</p> <p>Fase de treino: Modelo P: SE= 85,45% EP= 92,31% Modelo B: SE= 80,00% EP= 92,31% Especialistas: SE= 65,45% EP= 92,31%</p> <p>Fase de teste: Modelo P: SE= 82,35% EP= 91,78% Modelo B: SE= 79,41% EP= 91,78% Especialistas: SE= 79,41% EP= 86,30%</p> <p>Variáveis mais influentes: fluorescência, azul de toluidina, grau de infiltração, idade</p> <p>Seguimento de 3 anos: Modelo P previsão: 30 de alto risco e 71 de baixo risco Grupo de alto risco: 26.67% desenvolveram cancro Grupo de baixo risco: 7,04% desenvolveram cancro</p>	<p>Padronização permite reduzir a divergência devida a uma avaliação subjetiva</p> <p>Técnicas de deteção não invasivas: - Diferencia as lesões com elevada precisão - Fraca capacidade de prever o nível alto ou baixo do risco de cancro</p> <p>Modelo personalizado com dados pessoais: Modelo P apresenta os melhores resultados, inclusão de elementos pessoais adequados revelou-se benéfica para resolver a ambiguidade entre baixo e alto risco</p> <p>Classificação da influência das variáveis permitirá simplificar o modelo e aplicá-lo universalmente</p> <p>A previsão do modelo personalizado é válida a longo prazo, ultrapassa as previsões dos especialistas e pode prever o risco no futuro.</p> <p>Aplicação na internet em estudo</p> <p>Base de dados limitada</p>
--	--------------------------------------	---	------------------	--	--	---	---	--

CO : Cancro oral, DPMO : distúrbio potencialmente maligno oral, FA : Floresta aleatória

Tabela 7: Modelos de inteligência artificial com base de dados de fotografias orais

Autor/Ano/ País	Tipo de estudo	Amostra	Tipo de cancro	Método de classificação	Objetivo	Variáveis	Resultados	Conclusão
Fu Q, Chen Y, Li Z et al 2020 China	Estudo retrospectivo	5775 fotografias orais clínicas para dados de desenvolvimento e 401 fotografias para dados de validação interna de 11 hospitais na China entre 2006 e 2019. 402 fotografias clínicas de seis periódicos no campo da odontologia para dados de validação externa 666 fotografias orais do Hospital de Estomatologia da Universidade de Wuhan entre 4/11/2010 e 8/10/2019 para dados de validação clínica	CECO	- Redes Neurais Convolucionais (RNC) pré-treinados com ImageNet: DenseNet121 acoplado com <i>Single Shot MultiBox Detector</i> (SSD) - 21 leitores humanos (7 especialistas em cancro oral, 7 pós-graduados maxilofaciais, 7 não médicos) unicamente para análise de dados de validação clínica	Desenvolver uma abordagem de aprendizagem não invasiva, rentável e fácil de usar para identificar doentes com CECO utilizando imagens fotográficas.	Estádio de CECO, idade, localização das lesões, doenças da mucosa oral não-CECO, mucosa oral normal	Dados de validação interna: SE: 0.949, EP: 0.887, Precisão: 0.915 Análise secundária (validação interna): SE: 0.974, EP: 0.935, Precisão: 0.953 Dados de validação externa: SE: 0.896, EP: 0.806, Precisão: 0.841 Dados de validação clínica: - RNC: SE: 0.91, EP: 0.935, Precisão: 0.923 - Leitores humanos: Especialistas: SE: 0.917, EP: 0.931, Precisão: 0.924 Pós-graduados: SE: 0.831, EP: 0.907, Precisão: 0.87	Os resultados do RNC são muito próximos dos especialistas e excedem os dos estudantes. Pode-se argumentar que métodos de aprendizagem profunda podem oferecer possibilidades de identificação automática de doentes com CECO com desempenho equivalente ou mesmo superior ao de especialistas humanos treinados. Apesar das diferentes origens das fotografias, a DenseNet teve um bom desempenho e tem uma boa generalização que pode ser utilizada como uma ferramenta prática, não invasiva e rentável para os não especialistas detetarem lesões CECO. O estudo foi limitado pela insuficiente base de dados para treinar o algoritmo.

CECO : Carcinoma espinho-celular oral, RNC: rede neural convolucional, SSD: *Single Shot MultiBox Detector*

<p>Ferrer Sánchez A, Bagan J, Vila-Francés J et al.</p> <p>2022</p> <p>Espanha</p>	<p>Estudo diagnóstico comparativo</p>	<p>261 casos de leucoplasia oral com apoio de fotografias digitais standard do Serviço de Estomatologia e Cirurgia Maxilofacial do Hospital Geral Universitário de Valência.</p> <p>184 usados para treino e 77 para fase de teste</p> <p>Seguimento: 5,5 anos</p>	<p>CECO</p>	<p>Redes Neurais Convolucionais (RNC) supervisionado - Modelo tarefa individual -Modelo multitarefa (2 arquiteturas A e B)</p>	<p>Estimar a probabilidade de malignidade de uma lesão de leucoplasia oral usando o <i>Deep Learning</i> (DL), determinando o risco de desenvolver carcinoma espinocelular oral e o grau de displasia da lesão</p>	<p>3 variáveis: imagem, tipo clínico, localização</p>	<p>Transformação maligna: Sensibilidade: RNC individual: 1.00 RNC multitarefa: (A) 1.00 (B) 0.167 Especificidade: RNC individual: 0.692 RNC multitarefa: (A) 0.661 (B) 0.477 Precisão: RNC individual: 0.740 RNC multitarefa: (A) 0.714 (B) 0.428</p> <p>Displasia de alto risco: Sensibilidade: RNC individual: 0.857 RNC multitarefa: (A) 0.032 (B) 0.928 Especificidade: RNC individual: 0.740 RNC multitarefa: (A) 0.524 (B) 0.508 Precisão: RNC individual: 0.766 RNC multitarefa: (A) 0.441 (B) 0.584</p>	<p>Do ponto de vista clínico, é melhor que a taxa de falsos negativos esteja próxima de zero, portanto, é necessário maximizar a sensibilidade. As arquiteturas multitarefa são, portanto, melhores classificadores em comparação com o modelo individual. O modelo A tem melhor sensibilidade para transformação maligna e o modelo B para displasia. Assim, a melhor utilização possível deste modelo de RNC é explorar estes dois modelos e usar somente a saída mais prática de cada um. No entanto, o estudo tem limites, o número de participantes é reduzido e a classificação de displasia de alto risco é muito subjetiva de acordo com o patologista</p>
--	---------------------------------------	--	-------------	---	--	---	---	--

CECO : Carcinoma espinocelular oral, RNC: rede neural convolucional, DL: *deep learning*

<p>Warin K, Limprasert W, Suebnukarn S et al</p> <p>2022</p> <p>Thailand</p>	<p>Estudo comparativo retrospectivo</p>	<p>980 imagens fotográficas orais: 365 imagens de CECO, 315 imagens de DPMOs</p> <p>300 imagens não patológicas dos hospitais Oral and Maxillofacial Surgery Center of Thammasat University e Khon Kaen University</p>	<p>CECO, DPMOs</p>	<p>Modelos de classificação de imagens multiclasse: DenseNet-169, ResNet 101 pré-treinados com ImageNet, SqueezeNet, Swin- S</p> <p>10 especialistas (cirurgiões orais e maxilofaciais)</p> <p>10 médicos generalistas</p>	<p>Avaliar a eficácia dos algoritmos baseados na RNC na deteção e classificação de DPMO e CECO em imagens fotográficas orais e comparar com especialistas e generalistas.</p>	<p>Não especificadas</p>	<p>Dense-Net 169 obteve o melhor desempenho: CECO identificação: SE= 0.99, EP=0.99, ASC= 1.00, Precisão = 0.98</p> <p>DPMOs identificação: SE=0.95, EP=0.97, ASC=0.98</p> <p>Precisão = 0.95</p> <p>Dense-Net-169 e a ResNet-101 superaram mesmo o desempenho dos especialistas (CECO: SE= 0.9, EP= 0.89, DPMOs: SE= 0.74, EP=0.93)</p> <p>SqueezeNet e Swin-S superaram os generalistas e os seus resultados estão em concordância com os dos especialistas. Mas obtiveram resultados menos eficientes que RNC pesados (DenseNet169 e ResNet101)</p>	<p>Ferramenta de diagnóstico inovadora para ajudar os médicos, especialmente os clínicos gerais, a detetar lesões cancerosas com maior precisão e melhorar a deteção precoce em programas de rastreio do cancro oral.</p> <p>A desvantagem da Densenet e da ResNet é o seu tamanho e a necessidade de um processador de alto desempenho, que não pode ser facilmente aplicado.</p> <p>Pequena base de dados é uma limitação do estudo</p> <p>O objetivo é expandir o conjunto de dados de imagens para incluir outras lesões orais, desenvolvendo uma aplicação móvel baseada no RNC para recolher mais dados com um tamanho reduzido, como Squeeze-Net ou Swin-S.</p>
--	---	--	--------------------	--	---	--------------------------	---	--

CECO: Carcinoma espinocelular oral, DPMO : distúrbio potencialmente maligno oral, RNC: Rede neural conolucional, SE : sensibilidade, EP : especificidade, ASC : área sobre a curva, ResNet: *Residual Network*

Marzouk R, Alabdulkreem E, Dhahbi S et al. 2022 Arábia Saudita	Estudo diagnóstico comparativo	131 imagens de lábios e língua com malignidade (87 imagens) e sem malignidade (44 imagens).	CO	Modelo de deteção e classificação de cancro oral impulsionado por inteligência artificial com aprendizagem profunda de transferência (AIDTL-OCCM): Pré-processamento com fuzzy extração de características com DenseNet-169, classificação baseada em codificador automático e otimização de parâmetros baseada em Algoritmo de otimização Chimp - EfficientNet-b4 - Inception-v4 - DenseNet-161 - ResNet-152	Reconhecer o cancro oral utilizando a IA e técnicas de processamento de imagem	Não especificadas	AIDTL-OCCM: Precisão = 90.08% EfficientNet-b4 : Precisão =87.76% Inception-v4 : Precisão = 85.64% DenseNet-161 : Precisão =86.95% ResNet-152 : Precisão =87.45% O modelo AIDTL-OCCM alcançou o desempenho máximo em comparação com outros métodos.	Com a aplicação desses métodos de extração de dados de imagens, um modelo de IA mais eficiente foi desenvolvido para a deteção e classificação do CO.
--	--------------------------------	---	----	---	--	-------------------	--	---

CO: cancro oral, IA: inteligência artificial, ResNet: Residual Network

<p>Song B, Li S, Sunny S et al. 2021 India</p>	<p>Estudo diagnóstico comparativo</p>	<p>3851 imagens de luz branca polarizada da região jugal, 2417 normais, 1100 são casos pré-malignos, 243 são casos benignos e 91 são casos malignos das instituições Society Institute of Dental Sciences, Head and Neck Oncology Department of Mazumdar Shaw Medical Center e o Christian Institute of Health Sciences and Research.</p>	<p>CO</p>	<p>Modelo VGG19 pré-treinado com ImageNet</p> <p>-Abordagem ao nível dos dados: sobreamostragem e subamostragem ou associação de ambas</p> <p>-Abordagem ao nível do algoritmo: equilíbrio de pesos, novas funções de perda e aprendizagem de conjuntos</p>	<p>Aumentar o desempenho da IA utilizada para classificar imagens de CO utilizando conjuntos de dados multi-classe desequilibrados recolhidos em populações de alto risco</p>	<p>Não especificadas</p>	<p>Convencional pré-treinado VGG19: Precisão = 81%, mas classificação errada das classes minoritárias (casos benignos)</p> <p>-Abordagem ao nível do algoritmo: pequena melhoria, mas a maioria dos casos benignos continuava a ser classificada incorretamente</p> <p>- Abordagem ao nível dos dados: - sobreamostragem: desempenho do classificador de classes minoritárias aumentado - Associação com subamostragem (subamostragem das classes maioritárias, "normal" e "pré-malignidade", sobreamostragem das classes minoritárias, benignas e malignas): melhor desempenho do que utilizar apenas o aumento de dados</p> <p>-Associação de classificadores: reduz a variância e evita o sobre ajuste</p>	<p>Aplicando abordagens ao nível dos dados e ao nível do algoritmo a IA, é possível obter um bom desempenho em conjuntos de dados de imagens de cancro oral multi-classe desequilibrados.</p> <p>O desempenho das classes minoritárias, que inicialmente eram difíceis de distinguir, melhorou significativamente, embora a precisão global não tenha mudado muito. A precisão da classe de pré-malignidade também melhorou, tornando-a ótima para aplicações de rastreio.</p> <p>Precisa de ser testado noutras formas de aprendizagem profunda</p>
--	---------------------------------------	---	-----------	---	---	--------------------------	---	--

CO: Cancro oral, IA: inteligência artificial, VGGNet: *Visual Geometry Group Network*

Tabela 8: Modelos de inteligência artificial com base de dados adquirida com smartphone

Autor/Ano/ País	Tipo de estudo	Amostra	Tipo de cancro	Método de classificação	Objetivo	Variáveis	Resultados	Conclusão
Birur N P, Song B, Sunny S et al 2022 India	Estudo de observação prospetivo	Participantes elegíveis para análise: n = 4728 Participantes avaliados por FHW (Frontline-Health-Workers) n = 4728 Participantes avaliados por especialistas à distância: n = 4728 Participantes avaliados por especialistas no local: n = 752 Participantes avaliados por biopsia: n = 102 em hospitais com poucos recursos na Índia.	Cancro oral e DPMO	Dispositivo smartphone de imagem de modo duplo que combina Imagem de Luz Branca (ILB) e Imagem de Autofluorescência (IAF) associado: - Modelo MobileNet29 (pré-treinado com o conjunto de dados Imagenet) - Modelo RNC VGG19 Usado para FWH -Especialistas à distância -Especialistas no local	Verificar a precisão de um dispositivo de saúde móvel equipado com IA e RNC e usando ILB e IAF usado em conjunto com FHW para identificação de lesões orais suspeitas	Sexo, idade, fatores de risco (tabaco, areca), cor da lesão - Canais verde e vermelho do ILB e o rácio normalizado dos canais IAF-vermelho/verde	MobileNet integrada com o telemóvel: SE = 82%, EP = 77% Precisão = 79% VGG19 RNC integrado com o telemóvel: SE = 85% EP=82% Precisão = 85% Associação de VGG19 RNC e especialistas à distância em caso de incerteza: SE= 87% EP=86% Precisão = 87% Especialistas no local: SE= 94% EP= 72% Precisão =93% Especialistas à distância: SE=95% EP=84% Precisão = 92%	A sensibilidade do dispositivo tanto no telediagnóstico (95%) como no diagnóstico com MobileNet dos FWH (82%) foi equivalente ao exame direto por especialistas no local, demonstrando a viabilidade da técnica como um instrumento viável de despistagem do cancro oral em populações com recursos limitados e de alto risco. Telediagnóstico à distância por especialistas em casos de incerteza, assegurará a triagem precisa do sujeito e o encaminhamento adequado para um centro dedicado.

FWH : Frontline-health-workers, DPMO : Distúrbio potencialmente maligno oral, ILB: Imagem de Luz Branca, IAF: Imagem de Autofluorescência SE : sensibilidade, EP : especificidade, VGGNet: *Visual Geometry Group Network*

Jubair F, Al-karadsheh O, Malamos D et al 2021 Jordânia	Estudo diagnóstico comparativo	716 imagens clínicas para várias lesões de língua utilizando vários tipos de câmaras digitais e smartphones. - 480 imagens de lesões benignas - 236 imagens de lesões malignas ou DPMOs conjunto de treino (79%; n =566 imagens); conjunto de validação (7%; n =50 imagens); e conjunto de teste (14%; n =100 imagens).	Cancro da língua	Modelo de RNC pré-treinado com: - EfficientNet-B0 (peso leve, classificação binária), -VGG19, -ResNet101	Utilizando uma recolha de dados de imagens clínicas confirmadas, criar um RNC leve e profundo para discriminar entre lesões orais benignas e malignas ou potencialmente malignas.	Extraí automaticamente características discriminatórias examinando pixels de uma imagem policromática, sem necessidade de métodos artesanais de extração de características	EfficientNet-B0 : SE = 0.867, EP = 0.845 Precisão = 0.85 VGG19 : SE = 0.864, EP = 0.815 Precisão = 0.83 ResNet101 : SE = 0.839, EP = 0.844 Precisão = 0.84 A ASC para EfficientNet-B0 foi de 0,928 (95% CI: 0,88- 0,96), enquanto foi de 0,911 (95% CI: 0,87-0,95) para a VGG19 e 0,915 (95% CI: 0,88-0,95) para a ResNet101	Este RNC profundo associado com o modelo de transferência EfficientNet-B0 oferece grandes níveis de precisão, sensibilidade e especificidade para o diagnóstico de lesões orais cancerosas ou DPMOs, com os benefícios acrescidos de ser pequena em tamanho e usar pouco poder de processamento e memória. Esta técnica pode ser implementada como uma aplicação de smartphone, tornando-a um método potencialmente útil, acessível e de baixo custo. Para construir uma coleção substancial de fotografias clínicas para as DPMOs e o cancro oral, é necessária a cooperação internacional.
---	--------------------------------	--	------------------	---	---	---	---	--

DPMO : Distúrbio potencialmente maligno oral, RNC: Rede neural convolucional, SE : sensibilidade, EP : especificidade, ASC : área sobre a curva, VGGNet: *Visual Geometry Group Network*, ResNet : *Residual Network*

Lin H, Chen H, Weng L et al. 2021 China ¹⁰	Estudo retrospectivo	688 imagens de lesão oral e 760 imagens de mucosa normal Fase de treino: 993 imagens Fase de teste: 455 imagens	Cancro oral (CO)	3 passos RNC sistema : - capturar a lesão do doente com um smartphone com critérios específicos -Pré processamento - classificação da imagem com HRNet-W18 pré-treinado com ImageNet VGG16, ResNet50, DenseNet169	Apresentar um método eficaz de diagnóstico por imagem baseado em smartphones, alimentado por um algoritmo de aprendizagem profundo, para deteção automática de CO.	17M parâmetros	Melhor desempenho é obtido com posicionamento central e reamostragem (aumentar o conjunto de treino) HRNet-W18 alcançou SE de 83,0%, EP de 96,6% e precisão de 84,3%, ultrapassando outros IA, exceto DenseNet169 que tem EP quase similar. Taxas de erro relativamente mais altas para diferenciar lesões de baixo risco e alto risco. Baixa taxa de erro para deteção de tecido saudável e cancro oral (ASC respetivas 0.949 e 0.946)	O desempenho da deteção de cancro oral por IA pode ser melhorado com sucesso capturando fotos orais focadas na lesão, reamostrando os casos no conjunto de treino e aplicando o HRNet. Esta tecnologia com apoio de imagens de smartphones oferece uma promessa significativa para detetar o cancro oral em seus estádios iniciais. As imagens vêm de uma única fonte, é necessária mais diversidade de dados. O modelo pode ser útil para dentistas com pouca experiência, mas não pode tornar-se uma ferramenta para especialistas
---	----------------------	---	------------------	---	--	----------------	--	---

CO : cancro oral, RNC : rede neural convolucional, SE : sensibilidade, EP : especificidade, ASC : área sobre a curva, IA: inteligência artificial, VGGnet: *Visual Geometry Group Network*, ResNet : *Residual Network*, HRNet : *High-resolution Network*

Tabela 9: Modelos de inteligência artificial com base de dados amostras salivares

Autor/Ano/Pais	Tipo de estudo	Amostra	Tipo de cancro	Método de classificação	Objetivo	Variáveis	Resultados	Conclusão
Braz D, Neto M, Shimizu F et al. 2022 Brazil	Estudo diagnóstico comparativo	Amostras de saliva de 27 participantes dos quais 5 mulheres e 22 homens Controlo : n = 14 Cancro do pavimento bucal: n = 4 Cancro outra localização : n = 9	Cancro do pavimento bucal ou da cavidade oral	<i>Machine Learning</i> (ML): - supervisionado: Regressão Logística (RL), Análise Discriminatória Linear (ADL), K-vizinhos mais próximos (KNN), Naive Bayes (NB), e Máquina Vetorial de Suporte com núcleos: linear (MVSL), polinomial (MVSP) e função de base radial (MVSr), Floresta Aleatória (FA) associado ou não. - Espaço de Calibração Multidimensional (ECM)	Demonstrar que os dados de impedância obtidos com uma língua eletrônica em amostras de saliva podem ser utilizados para diagnosticar o cancro na boca com a ajuda de ML	23 variáveis (19 capacitâncias medidas em frequências que vão de 1 Hz a 1 MHz, idade, sexo, tabaco, alcoolismo) Com ECM : 26 variáveis (19 capacitâncias medidas em frequências que vão de 1 Hz a 1 MHz, idade, sexo, tabaco, alcoolismo)	- Para a classificação com apenas sensor, foram obtidos valores de alta precisão com MVSR e FA (Precisão semelhante considerando a dispersão). Precisão MVSR = 0.867(±0,208) (cancro classificação) Precisão FA = 0,630 (±0,052) (controlo classificação) - Para a classificação com todas as variáveis, FA é o mais eficiente e a inclusão de características clínicas proporcionou uma pequena melhoria no desempenho. Com associação ao ECM, obteve 80% de precisão média (com 65% de sensibilidade e 90% de especificidade). Variáveis mais importantes: frequência 215 Hz e " sem alcoolismo "	A combinação de FA e ECM reconhece potencialmente 65% dos doentes com cancro e 90% dos doentes saudáveis. A precisão tendia a aumentar quando a informação clínica dos doentes era utilizada em conjunto com os dados de impedância da língua eletrônica. Isto é encorajador para estudos futuros, uma vez que a combinação de dados de diferentes naturezas precisa de ser destacada no diagnóstico assistido por computador.

ML : *Machine learning*, RL : Regressão Logística, ADL: Análise Discriminatória Linear, KNN: K-vizinhos mais próximos, NB: Naive Bayes, e Máquina Vetorial de Suporte com núcleos: linear (MVSL), polinomial (MVSP) e radial (MVSr), FA: Floresta Aleatória, ECM: Espaço de Calibração Multidimensional

<p>Tseng Y, Wang Y, Hsueh P et al.</p> <p>2022</p> <p>Taiwan</p>	<p>Estudo retrospectivo</p>	<p>Amostras salivares de 337 doentes recrutados num hospital universitário Chi-Mei entre 09/2008 e 12/2012.</p> <p>80% para treino e 20% par teste</p>	<p>CECO</p>	<p>RL, FA, MVS, XGBoost, função de base radial (FBR) e um modelo de <i>staking</i> (combina as previsões de todos os modelos utilizados)</p> <p>ELISA</p>	<p>Desenvolver um modelo preditivo para identificar indivíduos com elevado risco de CECO utilizando IA que analisa biomarcadores de autoanticorpos salivares</p>	<p>10 proteínas relacionadas com o cancro oral, ANXA2, CA2, HSPA5, ISG15, KNG1, MMP1, MMP3, PRDX2, SPARC e p53 (intensidade de fluorescência num formato binário, logarítmico ou normalizado)</p> <p>Sexo, idade (num formato binário ou ternário), tabagismo, consumo de álcool e mastigação de noz de bétel</p>	<p>O método de <i>staking</i> foi determinado como o algoritmo de aprendizagem automática mais eficaz para prever casos de alto risco de CECO, com base nos seus resultados ASC nos conjuntos de teste.</p> <p>Melhor estratégia de processamento de dados: idade no formato original e níveis logarítmicos de autoanticorpos (ASC = 0,795 ± 0,055)</p> <p>Variáveis mais cotadas que contribuíram para a ASC: anti-ISG15, mastigação de noz de bétel, anti-ANXA2, anti-CA2 e anti-MMP3</p> <p>Utilização de marcadores autoanticorpos melhorou o desempenho de predição de 13,9%</p>	<p>Adição dos biomarcadores salivares melhora a precisão da previsão de risco para o CECO.</p> <p>Biomarcadores específicos para cada grupo (alto e baixo risco) foram identificados.</p> <p>4 das variáveis mais importantes para classificação dos dados foram biomarcadores.</p> <p>Com ajuda dos fatores de riscos tradicionais, uma abordagem viável é a utilização dos níveis de autoanticorpos salivares para o rastreio do CECO.</p> <p>Para validação externa, poderão ser necessárias investigações em vários centros.</p> <p>Diferentes biomarcadores devem ser investigados em grupos de estudo mais amplos (neste caso, pequena base de dados)</p>
--	-----------------------------	--	-------------	---	--	---	---	---

CECO: Carcinoma espinocelular oral, RL : Regressão Logística, FA: Floresta Aleatória, MVS : Máquina Vetorial de Suporte, FBR: função de base radial, ELISA: *enzyme-linked immunosorbent assay*, ASC : área sobre a curva

Tabela 10: Modelos de inteligência artificial com base de dados de técnicas de imagiologia

Autor/Ano/Pais	Tipo de estudo	Amostra	Tipo de cancro	Método de classificação	Objetivo	Variáveis	Resultados	Conclusão
Shao S, Mao N, Liu W et al 2020 China	Estudo retrospectivo	Imagens ponderadas por difusão de 269 participantes de janeiro 2015 até abril 2019 do Jining No. 1 People's Hospital 218 participantes com tumor benigno e 51 com tumor maligno 214 para fase de treino e 55 para fase de teste	Tumor das glândulas salivares (TGS)	Algoritmos de aprendizagem automática supervisionados: Regressão logística (RL), máquina de vetor de suporte (MVS) e K-vizinhos mais próximos (KNN)	Avaliar o sucesso da análise dos fatores radiômicos pelo classificador IA para identificar cancros epiteliais benignos e malignos das glândulas salivares em imagens ponderadas por difusão (DWI)	396 características radiômicas quantitativas: 42 histogramas, 9 fatores de forma, 10 Haralick, 11 matrizes de zonas de tamanho a nível de cinzento, 144 matrizes de co-ocorrência a nível de cinzento e 180 matrizes de comprimento de run99 8 selecionadas para LASSO (encolhimento mínimo absoluto e operador de seleção)	KNN apresentou uma ASC (0.791) mais baixa do que o RL(0.876) e o MVS(0.87). No conjunto de dados de teste, foi encontrado um resultado semelhante, ou seja, o valor ASC para RL foi de 0,876, seguido de MVS a 0,870 e KNN a 0,791. Fase de treino: RL: Precisão = 0.84 MVS: Precisão = 0.85 Fase de teste: RL: Precisão = 0.83 MVS: Precisão = 0.84	Os algoritmos RL e MVS apresentaram o melhor desempenho na classificação das lesões. Para diferenciar entre tumores benignos e malignos, as imagens DWI podem ser medidas utilizando fatores radiômicos operados por IA. Um limite do estudo é o tamanho da amostra. São necessárias populações de estudo maiores para confirmar os resultados Não foram incorporados no modelo aspetos morfológicos adicionais de RMN padrão, dados clínicos e genéticos. É ainda necessária mais investigação.

TGS : tumor das glândulas salivares, RL: regressão logística, MVS: máquina de vetor de suporte, KNN: K-vizinhos mais próximos, LASSO: *Least Absolute Shrinkage and Selection Operator*, ASC : área sobre a curva, DWI: *Diffusion-weighted Imaging*, IA: inteligência artificial, RMN: ressonância magnética nuclear

<p>Gunduz E, Alçin O, Kizilay A et al</p> <p>2022</p> <p>Turquia</p>	<p>Estudo retrospectivo</p>	<p>Imagem por ressonância magnética convencional e imagem ponderada por difusão (DWI) de glândulas parótidas de 123 participantes (grupo de controlo, grupo que foi submetido a parotidectomia entre 2010 e 2020)</p>	<p>Tumor das glândulas parótidas (TGP)</p>	<p>InceptionResNetv2, KNN, MVS (máquina vectorial de suporte), ADL</p> <p>Abordagens combinada (5 imagens de sequências RMN em conjunto) e não combinadas (cada sequência de RMN foi considerada um conjunto de dados de todas as variáveis</p> <p>Uso de votação por maioria (VM)</p> <p>Fase I : diferenciação dos tipos de tumores (grupo controlo, Tumor de Whartin, adenomas pleomórficos, tumor maligno)</p> <p>Fase II : diferenciação tumores benignos e malignos</p> <p>Fase III: diferenciação tumores e grupo controlo</p>	<p>Criar uma nova abordagem da IA baseada na aprendizagem profunda da RMN multiparamétrica no diagnóstico diferencial de tumores parotídeos comuns</p>	<p>T1Wce, T2W, b0 (DWI, b = 0 s/mm2), b1000 (DWI, b = 1000 s/mm2), e mapas de coeficiente de difusão aparente</p>	<p>Abordagem combinada teve o melhor desempenho em comparação com a abordagem não combinada O método VM aumentou a precisão em cerca de 10% e 6% nas abordagens não combinadas e combinadas, respetivamente para a fase I.</p> <p>Na fase II e III, o método MVS proporcionou o melhor desempenho. Além disso, o classificador InceptionResNetv2 ofereceu um desempenho exemplar no diagnóstico do TGP.</p> <p>Na fase I, o valor de precisão de InceptionRestNetv2 para a classificação nas abordagens não combinada e combinada era de 86,43% e 92,86%, respetivamente. Este valor na fase II e na fase III foi encontrado como 92,14% e 99,29%, respetivamente.</p>	<p>Esta abordagem combinada do DL com VM permite distinguir significativamente entre glândulas parótidas normais e cancerosas, bem como entre tumores benignos e malignos.</p> <p>A pequena base de dados é uma limitação do estudo para o <i>training</i> do algoritmo. Os modelos de aprendizagem profunda devem ser limitados ao apoio diagnóstico e associados às condições inflamatórias da glândula parótida ou envolvimento com doenças não tumorais para diagnóstico diferencial.</p>
--	-----------------------------	---	--	---	--	---	--	---

DWI : *Diffusion-weighted Imaging*, TGP: tumor da glândula parotídea, MVS: máquina de vetor de suporte, KNN: K-vizinhos mais próximos, ADL: Análise Discriminatória Linear, IA: inteligência artificial, RMN: ressonância magnética nuclear, VM: votação por maioria, MVS: máquina de vetor de suporte, DL: *deep learning*,

<p>He Z, Mao Y, Lu S et al 2022 China</p>	<p>Estudo diagnóstico comparativo</p>	<p>298 participantes foram inscritos neste estudo de Janeiro de 2009 a Setembro de 2020 no Hospital de Xiangya e Hunan Cancer Hospital, incluindo 56 doentes com tumor maligno, 108 com adenomas pleomórficos, 84 com tumor de Whartin, e 50 com outros tumores benignos.</p> <p>Imagens de RMN: Fase treino (70%), teste (30%)</p>	<p>Tumor da glândula parotídea (TGP)</p>	<p>- Aumento do gradiente extremo (XGBoost), MVS e árvore de decisão (AD) associados a um modelo de classificação binária em três etapas</p> <p>- 2 especialistas (5 e 10 anos de experiência em RM análise para deteção de cancro)</p>	<p>Avaliar o desempenho de modelos de aprendizagem de máquinas com base em características radiômicas de ressonância magnética morfológica na classificação de cancros parotídeos.</p>	<p>3396 características radiômicas: - 14 características baseadas na forma e no tamanho - 18 características de primeira ordem, baseadas em voxel único - Características de variações de nível cinzento - Características de primeira ordem e textura derivadas da transformação do filtro Wavelet - Características de primeira ordem e textura derivadas da transformação do filtro Laplacian de Gaussian</p>	<p>Primeiro passo (avaliando 6 características radiômicas): XGBoost superou MVS e AD (ASC = 0.826 contra 0.783 and 0.605) Segundo (12 características) e terceiro passo (8 características): MVS superou XGBoost (ASC= 0.833 contra 0.817 e ASC = 0.821 contra 0.789)</p> <p>Precisão : XGBoost (70,8%) superou o desempenho, seguido de MVS (59,6%), radiologistas (49,2%), e DT (46,1%)</p> <p>Para diferenciação dos tumores benignos e malignos, a precisão dos especialistas foi de 83,1% e de XGBoost de 89,9%.</p>	<p>IA parecia ter um desempenho superior ao dos radiologistas, e aceitável. O modelo de aprendizagem mecânica de três níveis era mais eficiente e preciso na classificação.</p> <p>Limitação porque não tem dados de DWI. Necessidade de realizar estudos com os outros parâmetros -(iômicos, genômicos, proteômicos..)</p> <p>IA pode ser usada na prática clínica como instrumento auxiliar de diagnóstico (primeiro rastreio e categorização dos TGP). Neste momento, os biomarcadores radiômicos só podem complementar o diagnóstico histológico.</p>
---	---------------------------------------	---	--	---	--	--	---	---

RMN: ressonância magnética nuclear, TGP : Tumor da glândula parotídea, MVS: máquina de vetor de suporte, AD: árvore de decisão, ASC : área sobre a curva, IA: inteligência artificial, DWI: *Diffusion-weighted Imaging*

Jeyaraj P, Samuel Nadar E 2019 India.	Estudo diagnóstico comparativo	100 imagens hiperespectrais das bases de dados do portal de BioGPS, <i>Cancer Imaging Archive</i> e <i>Genomic Data Commons</i>	CO	Rede neural de convolução (RNC) profunda particionada baseada na regressão MVS Rede de Crença Profunda (RCF)	Obter um desempenho comparável ao dos especialistas na classificação de imagens benignas e cancerosas e na identificação de regiões de interesse em imagens hiperespectrais multidimensionais	A intensidade da imagem, informação textual sobre informação espacial e espectral, a localização do voxel (pixel) e intensidade da imagem de vizinhança, as manchas cúbicas tridimensionais da imagem hiperespectral.	RNC particionada: SE= 0.91, EP=0.94, Precisão = 91.4% MVS: (Deepak Kumar et al.,2018) SE= 0.76, EP=0.86, Precisão = 82,4% RCF: (Dey et al., 2017) SE= 0.82, EP=0.89, Precisão = 84.5% Precisão aumentou de 4,5% utilizando um grande número de dados de indivíduos com cancro para a fase de treino (500 imagens) Os resultados classificados pelo RNC estão próximos dos resultados previstos pelos especialistas.	Este RNC profundo pode fornecer uma classificação precisa, constituir um classificador automático de imagens médicas sem conhecimentos de peritos e diferenciar entre tumores benignos e malignos. A utilização da rede neural de convolução particionada do sistema automático de diagnóstico do cancro é a mais adaptada para a análise de imagens hiperespectrais em comparação com os outros classificadores.
Aubreville M, Knipfer C, Oetter N et al 2017 Alemanha	Estudo Diagnóstico comparativo	7894 imagens de « Endomicroscopia confocal a laser (ECL)» de 12 doentes com cancro oral da área suspeita de malignidade (3815), de lábio inferior interno, o rebordo alveolar superior e a região do palato duro (4079 clinicamente normais)	CECO	Classificadores baseados em textura de imagem: Floresta Aleatória (FA) Redes Neurais Convolucionais (RNC) : Inception V3 pré-treinado com ImageNet ou com método de Fusão de Probabilidade por Patch	Apresentar e avaliar uma nova abordagem automática para diagnóstico do CECO utilizando tecnologias de aprendizagem profunda em imagens ECL. e comparar classificação baseada em características textuais e abordagens baseadas na RNC.	Classificação baseada em características textuais: -Padrões binários locais e matrizes de co-ocupação de nível cinzento RNC: sem variáveis pré-definidas, extração de características intrinsecamente dentro da rede.	Abordagem de fusão de probabilidade de patch (SE= 0.866, EP= 0.9, Precisão = 0,883, ASC= 0.955) tem um desempenho melhor do que os classificadores baseados em textura de imagem (SE= 0,847, EP= 0,782, Precisão = 0,814, ASC= 0,9) - Classificação baseada na aprendizagem por transferência de imagem quando se utiliza o RNC melhora o desempenho (SE= 0,9071, EP= 0,838, Precisão = 0,8702, ASC= 0,948.	Foi dado um grande passo com este estudo em direção à deteção automática de tumores malignos na imagem ECL. Com esta identificação em tempo real, poderá melhorar o processo clínico atualmente utilizado para o rastreio visual e táctil do cancro oral. Deve ser realizado um estudo mais aprofundado, incluindo um grupo de participantes saudáveis. Uma limitação da análise de imagem ECL é a diferenciação das DPMO e diferentes tumores benignos orais, só pode determinar se é benigno ou maligno

CO : Cancro oral, RNC: rede neural convolucional, MVS: máquina de vetor de suporte, RCF: Rede de Crença Profunda, SE : sensibilidade, EP : especificidade, ECL: Endomicroscopia confocal a laser, CECO: Carcinoma espinocelular oral, FA: Floresta Aleatória, ASC : área sobre a curva, DPMO : Distúrbio potencialmente maligno oral

<p>Heo J, Lim J, Lee H et al</p> <p>2022</p> <p>Coreia do Sul</p>	<p>Estudo diagnóstico comparativo</p>	<p>5576 imagens endoscópicas clínicas de cinco hospitais na Coreia do Sul entre 9 de Dezembro de 2010, e 24 de Setembro de 2020</p> <p>5224 imagens para dados de validação interna (treino)</p> <p>352 imagens para dados de validação externa (fase de teste)</p>	<p>Cancro da língua</p>	<p>RNC, ResNet, EfficientNet B3, VGGNet, MobileNet, DenseNet169</p> <p>2 leitores humanos (1 oncologista cirúrgico de cabeça e pescoço com mais de 7 anos de experiência clínica, 1 médico não-especialista com quatro anos de experiência)</p>	<p>Explorar a viabilidade de modelos de aprendizagem profunda baseados em endoscopia para o diagnóstico do cancro da língua com alta qualidade do conjunto de dados construído (dados clínicos).</p>	<p>Não especificadas</p>	<p>O modelo DenseNet169 mostrou o melhor desempenho. (ASC = 0.845, Área sob a curva Precision-Recall = 0.892)</p> <p>O modelo de aprendizagem profunda, médicos gerais e especialistas em oncologia tinham:</p> <ul style="list-style-type: none"> - Sensibilidades de 81,1%, 77,3% e 91,7%; - Especificidades de 86,8%, 75,0% e 90,9%; - Precisão s de 84,7%, 75,9% e 91,2%, respetivamente. <p>Foi demonstrada uma concordância justa entre o oncologista e o modelo desenvolvido para o diagnóstico do cancro (valor kappa = 0,685)</p>	<p>- Desenvolvimento dum algoritmo de aprendizagem profundo baseado na DenseNet169 com desempenho aceitável para o diagnóstico do cancro da língua a partir de imagens endoscópicas clínicas.</p> <p>Os modelos de IA têm o potencial de ajudar os médicos com pouca experiência clínica a diagnosticar precocemente o cancro oral, sem uma intervenção especializada.</p> <p>O âmbito do algoritmo é limitado, não pode distinguir entre diferentes lesões, tais como leucoplasias ou úlceras.</p> <p>Além disso, a aquisição de imagens precisa de ser padronizada (imagens de diferentes tamanhos e ângulos), o que pode enviesar os resultados. O diagnóstico do cancro não pode ser baseado apenas em imagens endoscópicas. Na investigação futura, espera-se que sejam desenvolvidos modelos de diagnóstico de alto desempenho se as imagens forem combinadas com vários dados clínicos</p>
---	---------------------------------------	---	-------------------------	---	--	--------------------------	---	---

RNC : rede neural convolucional, ASC : área sobre a curva, IA: inteligência artificial, ResNet: *Residual Network*, VGGNet: *Visual Geometry Group Network*

Tabela 11: Modelos de inteligência artificial com base de dados imagens de tomografia de coerência ótica (TCO)

Autor/Ano/Pais	Tipo de estudo	Amostra	Tipo de cancro	Método de classificação	Objetivo	Variáveis	Resultados	Conclusão
Yang Z, Shang J, Liu C et al 2022 China	Estudo diagnóstico comparativo	8680 imagens de tomografia de coerência ótica (TCO) obtidas de 14 doentes com um tumor das glândulas salivares (TGS) submetidos a tratamentos cirúrgicos no Hospital Estomatológico de Tianjin 5046 benignas (3961 adenoma pleomórfico (AP) e 1085 adenoma de células basais (ACB), e 3634 malignas (1677 carcinoma mucoepidermóide (CM) e 1957 carcinoma adenóide cístico (CAC) 80% para fase de treino e 20% para fase de teste	Tumor das glândulas salivares (TGS)	- RNA de retro propagação baseada na análise de componentes principais (PCA) - MVS	Desenvolver um algoritmo de classificação baseado em características de textura é utilizado para distinguir quatro tipos de tumores das glândulas salivares com base em imagens de tomografia de coerência ótica.	76 características de textura: - Histograma de níveis de cinzento, -Matriz de co-ocorrência de níveis de cinzento, -Abordagem espectral, -Medida de textura de Laws, -Auto-correlação centro-simétrica, - Padrão binário local Após a análise da variância explicada por cada componente, os primeiros cinco componentes contêm quase toda a informação importante	RNA de retro propagação: Precisão de todas as classes são superiores a 99% (Precisão média= 99,4%). - AP: SE=98,97%, EP=99,86%, - ACB: SE=99,52, EP=99,88%, - MC: SE= 99,67, EP=100%e - CAC: SE=99,60, EP=100% MVS: Precisão, sensibilidade e especificidade da RNA são superiores às da MVS e os resultados da classificação são mais estáveis	Devido às suas características de textura, as imagens TCO podem ser uma ótima ferramenta para diferenciar entre tumores benignos e malignos. Com base nos exames de TCO, é possível determinar as propriedades morfológicas das lesões e diferenciar entre estes 4 tipos de TGS. Para tornar este processo mais rápido e mais preciso, podemos utilizar PCA para reduzir a quantidade de dados que precisamos de analisar. Para aplicar a TCO ao ambiente oral é necessário desenvolver uma sonda de imagiologia portátil adequada. Para alargar a sua aplicabilidade às doenças orais, criação de mapas de características dos tecidos orais que sejam tão abrangentes quanto possível devem ser estudados.

TCO : tomografia de coerência ótica, TGS: tumor das glândulas salivares, AP : adenoma pleomórfico, ACB: adenoma de células basais, CM: carcinoma mucoepidermóide, CAC: carcinoma adenóide cístico, RNA: rede neural artificial, PCA: *Principal Component Analysis*, MVS: máquina de vetor de suporte

<p>Yuan W, Yang J, Yin B et al. 2022 China</p>	<p>Estudo diagnóstico comparativo</p>	<p>37 participantes com CECCO primário tratados de Setembro de 2020 a Junho de 2021 no Hospital Estomatológico de Pequim 468 imagens 3D, (224 dados malignos e 244 dados benignos)</p>	<p>CECCO</p>	<p>Aprendizagem residual profunda multinível (ARPM) desenvolvido a partir da rede neural residual (ResNet) -VGGnet -GoogleNet -ResNet -Especialistas</p>	<p>Criar um ARPM para identificar tecidos malignos e benignos a partir de imagens TCO (tomografia de coerência ótica), encontrando os parâmetros que minimizam a diferença entre os resultados esperados e previstos do classificador</p>	<p>- Classificação piramidal das características de alto e baixo nível graças aos blocos residuais das camadas convolucionais precedentes - Mecanismo de atenção espacial na última característica piramidal</p>	<p>Classificação ao nível da imagem: sistema ARPM atinge uma SE de 91,2%, uma EP de 83,6% e uma precisão de 87,5%, é o melhor desempenho de diagnóstico entre os diferentes métodos. VGG, ResNet e GoogLeNet atingem ASCs de 0,868, 0,895 e 0,917, respectivamente, enquanto o ARPM atinge a maior ASC de 0,920. Classificação ao nível das "Patches" ou secções: O diagnóstico é mais preciso do que ao nível da imagem, melhoria para cada método RNC ARPM: SE=100%, EP= 86,7%, precisão = 93,1% ARPM supera os especialistas na EP (de 17,4%), precisão (de 6.9%) Com sensibilidade equivalente</p>	<p>Ao incluir características de vários níveis e um mecanismo de atenção espacial, a ARPM supera significativamente as RNCs existentes e tem potencial para o diagnóstico não invasivo do CECCO No futuro, o maior obstáculo será a disponibilidade de um número suficiente de imagens TCO anotadas para treinar o modelo. Vai ser necessário estudar a aprendizagem semi-supervisionada e não-supervisionada.</p>
--	---------------------------------------	--	--------------	--	---	---	---	---

CECCO: Carcinoma espinocelular oral, ARPM: Aprendizagem residual profunda multinível, TCO: tomografia de coerência ótica, SE : sensibilidade, EP : especificidade, RNC : rede neural convolucional, ASC : área sobre a curva; VGGNet: *Visual Geometry Group Network*, ResNet : *Residual Network*

Tabela 12: Modelos de inteligência artificial com base de dados de análise de biomarcadores

Autor/Ano/Pais	Tipo de estudo	Amostra	Tipo de cancro	Método de classificação	Objetivo	Variáveis	Resultados	Conclusão
Hsieh C, Chen W, Hsieh Y et al 2018 Taiwan	Estudo diagnóstico retrospectivo	Foram recrutados 38 participantes com CECCO primários não tratados e 84 sujeitos saudáveis para o estudo no Hospital Memorial Chang Gung de Agosto de 2015 a Março de 2017 134 miRNAs selecionados isolados de amostras de sangue analisados com PanelChip™ Analysis System, PCR 90% para treinar e 10% para fase de teste	CECCO	MVS (máquina vetorial de suporte)	Desenvolver um método de deteção do CECCO baseado na análise da expressão miRNA utilizando inteligência artificial a partir de amostras de sangue	ΔCq (variação do número de ciclos de PCR)	MVS: SE = 0.816 EP = 0.988 o classificador é adequado para identificar indivíduos saudáveis devido à sua elevada especificidade e tem uma sensibilidade aceitável. Número médio de miRNAs, de 134 do classificador, detetado no plasma: Sujeitos saudáveis = 93 Sujeitos com CECCO = 97	Este estudo permite descobrir um padrão de expressão miRNA, utilizando a combinação do sistema PanelChip™ Analysis System Chips e o classificador de CECCO, capaz de separar sujeitos saudáveis e doentes com CECCO Será necessário realizar o estudo com um maior número de indivíduos. Os fatores demográficos e clinicopatológicos dos doentes inscritos não foram tidos em conta. Pode ser feito outro estudo com estes fatores para ver se melhora o desempenho da IA

CECCO: Carcinoma espinocelular oral, RNA: ácido ribonucleico, MVS: máquina vetorial de suporte, PCR: *Polymerase Chain Reaction*, SE : sensibilidade, EP : especificidade, IA: inteligência artificial

<p>McRae M, Modak S, Simmons G et al.</p> <p>2020</p> <p>Estados Unidos</p>	<p>Estudo prospetivo</p>	<p>Amostras histopatológicas e citológicas de escova de 714 sujeitos de 3 grupos de participantes:</p> <ul style="list-style-type: none"> - com DPMOs diagnosticados por biopsia - com lesão maligna recentemente diagnosticada - indivíduos saudáveis 	<p>CECO</p>	<p>KNN e modelo de análise de regressão logística baseado no software CellProfiler</p> <p>Com análise de componentes principais para simplificar o conjunto de dados complexo (PCA)</p>	<p>Demonstrar como uma plataforma de citologia compatível com ponto de atendimento pode ser útil para rastrear e monitorizar lesões orais e distinguir todo o espectro da displasia epitelial oral.</p>	<p>- Fenótipos celulares (ML modelo): 144 características celulares/nucleares de análises unicelulares, incluindo medições morfológicas e baseadas em intensidade.</p> <p>6 biomarcadores moleculares ($\alpha\beta6$, EGFR, CD147, McM2, gemininin, e Ki-67)</p> <p>-DPMOs : distribuições de fenótipo celular , idade, sexo, tabagismo, tamanho da lesão, impressão clínica da lesão e cor da lesão</p>	<p>96.9% de precisão para distribuição de fenótipos celulares</p> <p>Lesão benigna: tipo 1 (maduras)</p> <p>Lesão maligna: tipo 2 (pequenas), tipo 3 (leucócitos), e tipo 4 (núcleos solitários).</p> <p>-Benigna contra displasia: SE= 0.69, EP = 0.77, ASC= 0.81</p> <p>-Displasia leve versus moderada: SE=0.79, EP= 0.85, ASC= 0.88</p> <p>-Baixo contra alto risco de malignidade: SE= 0.78, EP= 0.87, ASC= 0.88</p> <p>-Displasia benigna contra maligna: SE=0.89, EP=0.9, ASC=0.95</p> <p>-Controlo saudável contra maligno: SE= 0.94, EP= 0.92, ASC= 0.97</p>	<p>As alterações fenotípicas celulares observadas nas amostras citológicas estão correlacionadas com o nível de doença e podem assim ser úteis na avaliação de DPMOs.</p> <p>Este software simples e sensível de citopatologia pode fornecer uma triagem não invasiva para DPMOs detetados e pode ajudar patologistas especializados e não especializados. O trabalho futuro pode expandir a utilização de biomarcadores moleculares e explorar a identificação de fenótipos celulares raros adicionais para melhorar ainda mais o desempenho e diferenciação entre as displasias (diferenciar entre cãndida e leucoplasia,..)</p>
---	--------------------------	---	-------------	---	---	--	---	--

DPMO : Distúrbio potencialmente maligno oral, CECO: Carcinoma espinocelular oral, KNN: K-vizinhos mais próximos, PCA: *Principal Component Analysis*, ML: *machine learning*, SE : sensibilidade, EP : especificidade, ASC : área sobre a curva

<p>Mentel S, Gallo K, Wagendorf O, et al.</p> <p>2021</p> <p>Alemanha</p>	<p>Estudo controlado prospetivo</p>	<p>Amostras de respiração de 55 participantes com potencial CECO antes e após cirurgia, bem como 50 amostras de respiração de controlos saudáveis fumadores realizadas com BreathSpec</p>	<p>CECO</p>	<p>FA, KNN, RL, análise discriminante linear (ADL)</p>	<p>Avaliar a possibilidade de utilizar o teste respiratório como método de deteção do cancro em doentes com CECO</p>	<p>Compostos orgânicos voláteis específicos analisados com cromatografia gasosa e espectrometria de massa de iões (diferença de intensidade com cores vermelho, verde, azul)</p>	<p>O estudo foi capaz de distinguir amostras de tumores pré-operatórios de amostras de fumadores saudáveis com uma precisão média de 0,90(desvio positivo) e 0,86 (desvio negativo)</p> <p>A distinção entre amostras pré e pós-operatórias só pôde ser efetuada com uma precisão média máxima de 0,65(desvio positivo) e 0.89 (desvio negativo).</p> <p>RL foi considerado o método com melhor desempenho global.</p>	<p>Uma comparação entre doentes com CECO e fumadores saudáveis mostrou que algumas substâncias eram mais prevalentes em doentes com CECO.</p> <p>Os modelos foram capazes de distinguir entre doentes com CECO e voluntários saudáveis</p> <p>Mas nesta fase, ainda não é possível distinguir doentes com elevado risco de cancro oral. São necessários mais estudos, centrados no reconhecimento de padrões utilizando a IA numa amostra maior e em estudos <i>in vitro</i> em tecidos tumorais para descobrir compostos orgânicos voláteis específicos.</p>
---	-------------------------------------	---	-------------	--	--	--	--	---

CECO: Carcinoma espinocelular oral, FA: Floresta aleatória, KNN: K-vizinhos mais próximos, RL: Regressão logística, ADL: Análise discriminante linear, IA: inteligência artificial

5. Discussão

Dado que a inteligência artificial está a ser cada vez mais desenvolvida no sector médico, o nosso estudo procurou demonstrar se esta poderia ser aplicada à deteção precoce do cancro oral. Para podermos ter uma imagem do potencial desta aplicação, analisámos os diferentes métodos de deteção do cancro oral e a sua associação com os diferentes tipos de IA mencionados na nossa introdução teórica.

5.1. Principais IAs utilizadas e suas vantagens

Um dos tipos de exploração de IA mais utilizados nos vários estudos é a rede neural convolucional (RNC) profunda. De facto, mais de metade dos estudos (3,10,15–18,21,22,25,28,30–32) utilizam a aprendizagem profunda como base para o seu modelo de classificação. Em vez de utilizarem a extração de características como primeiro passo, as redes neuronais convolucionais (RNC) aceitam uma imagem como entrada e, em seguida, efetuam a extração de características. (21) A existência de camadas específicas de conjugação e convolução, que lhes conferem a capacidade de extrair estas características significativas e reconhecer padrões complexos, faz das RNC profundas uma estrutura altamente eficaz para o reconhecimento de imagens. (18) ResNet(3,15,17,18,21,25,31), DenseNet (3,16,18,21,31) e VGGNet(10,15,17,21,30,31) são os modelos de aprendizagem profunda mais frequentemente testados nos nossos artigos. Estas redes neuronais são exploradas pelo seu esqueleto altamente desenvolvido. À medida que as camadas se multiplicam e aprofundam, estes modelos podem alcançar um melhor desempenho de previsão. (21)

No entanto, à medida que o sistema se aprofunda, a camada anterior da rede é recuperada e os dados de entrada ou o gradiente que passou pela camada desvanecem-se gradualmente. Assim, ao ligar instantaneamente todas as camadas com tamanhos de características comparáveis, a DenseNet resolve o problema do desvanecimento dos gradientes. A principal razão para utilização do modelo DenseNet como um extrator de características é que, ao aprofundar a rede, é possível encontrar características mais gerais, permitindo um melhor desempenho. (18)

Contudo, as redes pesadas como a ResNet e a DenseNet têm um tempo de aprendizagem muito lento em comparação com os modelos VGGNet ou MobileNet, que são mais leves e, por conseguinte, mais rápidos. (18) Warin K. et al compararam o desempenho de RNCs pesadas, DenseNet-169 e ResNet-101, e de RNCs leves, SqueezeNet e Swin-S, para a classificação de imagens de CO e DPMOs. Os modelos mais leves obtiveram resultados promissores com uma área sob a curva entre 0,71 e 0,88, mas tiveram um desempenho inferior ao das RNCs mais pesadas.

No entanto, apesar do seu melhor desempenho, a DenseNet e a ResNet, devido à necessidade de servidores potentes, não podem ser aplicadas em telemóveis, por exemplo, num consultório dentário.(3)

A escolha do modelo depende, portanto, do desempenho esperado, do tipo de base de dados analisada para a aprendizagem, da velocidade desejada e do contexto em que o modelo vai ser utilizado.

Exigindo uma quantidade muito menor de dados de treino (22), os modelos de classificação clássicos também são promissores para a deteção precoce do cancro oral. De facto, 63% dos artigos realizaram testes de diagnóstico utilizando modelos clássicos de aprendizagem automática. (7,19,20,22–27,29,32–34). A floresta aleatória (FA) (20,22,26,29,33) e Máquina Vetorial de suporte (MVS)(19,20,23–25,27,32,33) são os modelos clássicos de aprendizagem automática mais frequentemente mencionados nos estudos.

Devido ao seu desempenho notável, o FA é amplamente utilizado no domínio da IA. (35)O seu funcionamento consiste em calcular a média das previsões de várias árvores de decisão, cada uma das quais lida com uma categoria diferente de fatores, de modo que todos os fatores a ter em conta possam ser devidamente agrupados. (26)

O MVS, por outro lado, tem uma forte capacidade de generalização, ou seja, a capacidade de aplicar os conhecimentos adquiridos durante a formação para fazer previsões exatas sobre dados que nunca viu antes, nomeadamente para uma amostra de dimensão limitada, graças à sua noção de minimização do risco estrutural. Outra vantagem da MVS é a sua sólida base teórica matemática, que lhe permite encontrar uma solução ótima para um determinado problema. Por conseguinte, as MVS foram utilizadas em vários estudos como um método de aprendizagem automática encorajador.(24)

O estudo de Braz et al. comparou o desempenho destes dois algoritmos de ML. Cada um dos dois modelos de aprendizagem obteve os melhores resultados de classificação de

acordo com parâmetros distintos, MVS para reconhecimento ou não reconhecimento de casos de cancro e FA para reconhecimento de casos de controlo. (33) Para o explorar da melhor forma, o quadro e o objetivo da aplicação da IA não devem ser negligenciados. Por conseguinte, é essencial utilizar cada modelo de IA no contexto em que terá o melhor desempenho.

Tseng et al. avaliaram o MVS e o FA e concluíram que, ao combinar os resultados dos algoritmos utilizando o método de *stacking*, foi possível obter um melhor desempenho na deteção do cancro oral.(20)

Por conseguinte, podem ser desenvolvidas diferentes abordagens, pelo que a escolha adequada do algoritmo é vital para a deteção precoce do CO mais adequado, como a escolha de um determinado modelo pelo seu desempenho numa dada situação ou a combinação de várias IA.

A mesma observação foi feita por Ferrer et al. relativamente à aprendizagem *deep learning*, comparando duas arquiteturas diferentes de modelos multitarefa. Após a análise dos resultados, sugerem a utilização dos 2 modelos e a escolha do melhor rendimento em função do objetivo pretendido, a deteção da transformação maligna ou a deteção da displasia. (28)

Para saber qual o tipo de IA a selecionar, *deep* ou *machine learning* clássico, na análise de imagens hiperespectrais, Jeyaraj et al. compararam os dois tipos com MVS e uma RNC profunda. A análise do desempenho levou à suposição de que a RNC profunda é mais competente para explorar dados hiperespectrais, devido à complexidade dos fatores a ter em conta. (32) A quantidade e a variabilidade dos dados também devem ser consideradas ao escolher uma IA.

No caso da deteção precoce do cancro oral, ambas as abordagens de aprendizagem podem, por conseguinte, ser utilizadas com êxito, mas a escolha dependerá das especificidades da tarefa, da quantidade e qualidade dos dados disponíveis bem como dos recursos informáticos e do tempo necessário para a formação e aplicação do modelo.

Outros métodos de otimização da aprendizagem também foram explorados através do *transfert learning*, que recorre ao conhecimento prévio de um modelo pré-treinado que foi aplicado num conjunto de grandes bases de dados de imagens para melhorar o desempenho do modelo utilizado no estudo, de acordo com Fu et al.(16), permitindo um ganho de tempo e um melhor desempenho quando os dados disponíveis para o estudo são limitados.

Por exemplo, Jubair et al. provaram que a qualidade dos dados de aprendizagem por transferência tinha precedência sobre a dimensão dos dados através à utilização da EfficientNet-B0, o que lhes permitiu obter um desempenho na deteção de CO superior ao associado ao treino do modelo através de IAs pesadas. A quantidade de dados aproveitáveis para treino é, portanto, maior para a EfficientNet-B0, apesar da menor quantidade disponível.(15)

Vários autores procuraram demonstrar se todas estas inovações e vantagens proporcionadas pela IA igualam ou ultrapassam o desempenho de dentistas e especialistas. Os resultados obtidos por Warin et al, Heo et al e Fu et al parecem concordar entre si, com a análise dos modelos de aprendizagem a exceder a dos profissionais de saúde sem experiência.(3,16,21) IA pode, por conseguinte, ser utilizada como uma ferramenta de rastreio e ajudar os profissionais a melhorar a sua precisão de diagnóstico. No entanto, a comparação com os especialistas é mais heterogénea, mesmo que os algoritmos pareçam ser mais eficazes. De acordo com Heo et al, o Densenet169 tem um desempenho 10% pior do que um especialista e tem dificuldade em distinguir entre diferentes lesões de DPMOs. A opinião do especialista é, por conseguinte, crucial para a deteção de lesões de risco, apesar dos resultados aceitáveis do DenseNet169 para a deteção de CO, que apenas pode ser considerado como um auxiliar.(21) Embora os seus resultados sugiram que os modelos superam os especialistas, Warin et al. e He et al. concordam com Heo et al. que a IA como auxiliar no diagnóstico do cancro oral não substitui a competência e a experiência clínica dos médicos dentistas, mas fornece uma ferramenta adicional para melhorar a precisão do diagnóstico e facilitar a tomada de decisões clínicas.(3,27) Apenas Wang et al. e Fu et al. propõem que a deteção automática do cancro possa ser encarada no futuro sem a necessidade de intervenção de um especialista. (16,26)

Após ter aprofundado as diferentes formas de IA, é necessário analisar os diferentes métodos de aquisição de dados e o seu impacto na deteção precoce do cancro oral para compreender toda a extensão da aplicação da IA neste domínio.

5.2. Métodos de aquisição de dados

6.2.A. Aplicações clínicas e hospitalares

A imagiologia, um dos métodos mais comumente adotados para a deteção do cancro oral, tem sido estudada em vários artigos como uma base de dados promissora. Em particular, a ressonância magnética (RM) tem sido objeto de muita investigação. Gunduz et al. apresentaram um estudo que demonstra a eficácia de um novo método que combina imagens de RMN e de DWI, ponderadas por difusão, diretamente fornecidas em conjunto como entrada para o modelo de aprendizagem, InceptionResNetv2, combinado com a votação por maioria após a análise de cada imagem. A utilização do *deep learning* combinado com a votação por maioria permite determinar uma decisão final tendo em conta as previsões de cada imagem. Esta abordagem melhorou a fiabilidade e a precisão do modelo de deteção de CO e de diferenciação entre diferentes tipos de tumores, explorando a diversidade de informações fornecidas por cada imagem. (25).

He et al. adotaram outra abordagem para a exploração dos dados de RMN, incluindo a radiômica. A radiômica pode ser utilizada para extrair características quantitativas, como a forma, a textura, a intensidade e as relações espaciais de regiões de interesse em imagens que possam estar relacionadas com lesões cancerígenas e, por conseguinte, com a sua deteção. (27) Com um tamanho de amostra muito superior ao da pesquisa de Gunduz et al., o estudo é mais representativo da realidade, proporcionando maior fiabilidade. No entanto, o estudo de Gunduz et al. obteve melhores resultados na deteção de tumores da parótida e na sua classificação diferencial. (25) Essa diferença de desempenho pode ser devida ao tamanho da amostra e à sua composição. Quando a amostra é menor, pode haver maior variabilidade e maior sensibilidade às características específicas desse grupo restrito.

Combinando imagens de RMN, DWI e características radiômicas, Shao et al. tentaram desenvolver algoritmos de aprendizagem automática capazes de identificar tumores

epiteliais benignos e malignos. Os fatores radiológicos foram selecionados por LASSO, tal como no estudo de He et al. mas a sua análise não foi faseada, podendo esta diferença na exploração dos resultados ser uma das razões para o desempenho inferior alcançado por Shao et al. A utilização do método LASSO para selecionar os fatores radiômicos pode ser considerada uma abordagem promissora, uma vez que permite reduzir a dimensionalidade dos dados através da identificação das características mais relevantes. No entanto, a forma de exploração e análise desses fatores radiômicos diferiu entre os dois estudos, o que pode ter tido consequências no desempenho obtido. O artigo de Shao et al, que é mais antigo, utiliza apenas a aprendizagem automática clássica, em comparação com a RNC de Heo et al, que pode ser mais adequada para explorar estes dados complexos. (23,27)

Importa salientar que os três estudos concordam que a IA deve continuar a ser uma ajuda auxiliar no diagnóstico e diferenciação das lesões. Embora a IA possa ser uma ferramenta poderosa para apoiar os dentistas na deteção do cancro oral, não deve substituir os conhecimentos e a avaliação clínica dos profissionais de saúde. (23,25,27)

Outros trabalhos abordaram técnicas de imagiologia que oferecem uma maior resolução para visualizar e caracterizar tecidos biológicos a nível microscópico ou subcelular. Jeyaraj et al. exploraram dados de imagens hiperespectrais, oferecendo uma maior precisão graças aos dados espaciais, espectrais e texturais. Os resultados da investigação levaram à conclusão de que a utilização de uma rede neural particionada é o método mais adequado para explorar estes dados. (32)

O estudo de Aubreville et al. sobre as imagens ECL "*Endomicroscopia Confocal a Laser*", que têm uma boa penetração na avaliação dos tecidos, abriu caminho a uma exploração muito prometedora destes dados. Os investigadores conseguiram tirar partido das informações fornecidas pelas imagens ECL para obter resultados significativos, extraíndo os dados sob a forma de *patches*. (22) No entanto, estes dois métodos apenas identificaram lesões malignas e benignas, sem qualquer diferenciação entre os diferentes tipos de lesões. (22,32)

Os trabalhos de Yang et al. e Yuan et al. também destacaram a utilização da tomografia de coerência ótica (TCO) como um método de imagem alternativo para o rastreio do cancro, conhecido pela sua alta resolução, que fornece uma grande quantidade de informações sobre as texturas das imagens. (17,24) O artigo de Yang et al. apresenta-nos um avanço neste domínio com a diferenciação de 4 tipos de tumores da parótida utilizando a PCA "*Principal Component Analysis*", ao contrário dos outros dois métodos apresentados anteriormente, que não eram capazes de os distinguir. A PCA permitiu simplificar a representação e a utilização dos dados, reduzindo a sua dimensionalidade, e assim aumentar o desempenho da IA. (24) Yuan et al. combinaram o *transfert learning* com a extração de dados sob a forma de *patches*, o que parece ser uma das soluções mais adequadas para lidar com a complexidade das imagens de TCO. (17) A utilização de dados TCO pela IA parece ter potencial futuro na deteção precoce do cancro, sendo os resultados dos estudos de Yuan et al. e Yang et al. muito conclusivos, com taxas de sucesso próximas de 100%, abrindo assim novas perspetivas de rastreio. (17,24)

Combinando um algoritmo de aprendizagem adaptado com uma extração e um processamento de dados otimizados, é possível fazer progressos significativos na deteção precoce do cancro oral. Tal permitirá melhorar a sensibilidade e a especificidade dos métodos de diagnóstico, identificar com maior precisão as lesões suspeitas e facilitar a tomada de decisões clínicas.

Como a imagiologia não é o único meio de deteção precoce do cancro, foram desenvolvidos outros meios de aquisição, numa tentativa de tornar o exame do doente o menos invasivo possível. O estudo dos biomarcadores também se tem destacado no diagnóstico precoce do cancro oral, nomeadamente os biomarcadores salivares. Braz et al. e Tseng et al. debruçaram-se sobre este tema e obtiveram resultados bastante conclusivos graças à utilização de fatores totalmente diferentes como entrada para o algoritmo. O artigo de Braz et al. centra-se na utilização de uma língua eletrónica para a discriminação de amostras de saliva, enquanto o artigo de Tseng et al. se centra na intensidade de fluorescência de biomarcadores de autoanticorpos salivares para a previsão do risco de carcinoma de células escamosas da cavidade oral. Ambos concordam com a dificuldade de adaptar os dados para melhorar o desempenho do algoritmo. (20,33)

Braz et al., graças a estes dados resultantes da interação entre as amostras biológicas e os sensores da língua eletrónica, podem diferenciar as amostras de doentes com cancro das amostras de pessoas saudáveis, sem necessidade de conhecer especificamente o biomarcador ligado ao cancro. (33) O método de Tseng et al., por outro lado, requer a deteção de biomarcadores. O algoritmo de aprendizagem utilizado no estudo permitiu identificar determinados marcadores biológicos específicos do cancro oral e confirmou que estes eram significativamente mais elevados no grupo de alto risco do que no grupo de baixo risco. De fato, a IA tem a capacidade de fazer comparações pormenorizadas e de procurar padrões específicos para identificar os detalhes que podem passar despercebidos ao olho humano.

Esta sensibilidade da IA pode permitir a descoberta de novos biomarcadores de cancro potencialmente relevantes, ainda não identificados, e ter um impacto significativo no diagnóstico precoce e, conseqüentemente, no tratamento do cancro oral. (20)

Mentel et al. analisaram amostras respiratórias de doentes, outro marcador biológico, e recorreram à análise de IA para identificar a presença de compostos voláteis presentes em grandes quantidades apenas em indivíduos saudáveis. (29) Hsieh et al. recorreram a dados de miRNA para detetar uma maior quantidade de miRNA em doentes com CO do que em indivíduos saudáveis, utilizando o seu sistema miRScan™ PanCancer Chips em conjunto com o modelo de aprendizagem, e conseguiram identificar um perfil de expressão de miRNA que permitiu diferenciar os dois grupos. (19) Os resultados destas investigações ainda não são conclusivos e é necessário efetuar mais estudos, mas o interesse pelo estudo dos biomarcadores através da IA é muito promissor para o futuro.

De facto, a descoberta de novos biomarcadores graças à IA pode abrir novas perspetivas no domínio da investigação sobre o cancro. Estes biomarcadores podem fornecer informações valiosas sobre os mecanismos biológicos subjacentes ao cancro, o que pode levar a uma melhor compreensão da doença e a novas estratégias terapêuticas.

No entanto, todas estas aplicações estão limitadas ao domínio clínico. Por serem dispendiosos e exigirem ferramentas específicas, é difícil implementá-los no primeiro ponto

de atendimento do doente, o consultório dentário. Como o objetivo é detetar o cancro precocemente, estes métodos só podem ser aplicados após o rastreio inicial, motivo pelo qual foram também estudados outros métodos de mais fácil acesso e aplicação.

6.2.B. Aplicáveis a práticas ou áreas com poucos recursos

Mais facilmente exploráveis no consultório, as investigações recentes exploraram a utilização da IA para detetar o cancro oral utilizando métodos acessíveis a qualquer profissional de saúde. Alhazmi et al. desenvolveram um modelo de aprendizagem de rede neural, com resultados conclusivos, fazendo previsões baseadas apenas nas informações clínicas dos doentes. Esta abordagem pode ser facilmente integrada pelo dentista no consultório e utilizada para o rastreio preliminar, ajudando o dentista no seu plano de tratamento, por exemplo, através do desenvolvimento de uma plataforma Internet adequada.(34)

Noutro método fácil de aplicação, Ferrer et al. recolheram imagens de cavidades orais de um departamento hospitalar para serem analisadas por uma IA que combina segmentação supervisionada com uma rede neural convolucional. (28)Warin et al. utilizam o mesmo tipo de dados, fotografias orais clínicas, mas desta vez comparando várias IAs supervisionadas multiclasse, incluindo a DenseNet169 pré-treinada por aprendizagem por transferência com a ImageNet e com anotações de 3 especialistas diferentes para o treino.(3) Ferrer et al. obtiveram uma sensibilidade, especificidade e precisão quase iguais a 1, tornando-o o estudo com melhor desempenho para a deteção precoce do cancro oral. No entanto, os resultados de Warin et al. são muito semelhantes e esta diferença pode dever-se unicamente ao tamanho da amostra e à sua distribuição, sendo a amostra de Warin et al. quase 4 vezes maior do que a de Ferrer. Ambos os estudos concordam que os modelos mais complexos, o multitarefa, comparado com a tarefa individualizada, no estudo de Ferrer et al. e o modelo pesado no estudo de Warin et al. são os mais adequados para analisar os dados fotográficos. (3,28)

Apesar dos seus excelentes resultados, estes modelos continuam a exigir a intervenção humana para anotar as imagens na fase de pré-processamento, o que pode exigir muito tempo e recursos para recolher, anotar e preparar os dados. O modelo de aprendizagem supervisionada pode também registar uma quebra de desempenho quando confrontado com situações novas ou dados não rotulados.

Marzouk et al. propuseram-se resolver este problema e tentaram desenvolver o seu próprio modelo não supervisionado. Graças à sua metodologia adaptada de pré-processamento de dados, utilizando a extração de dados difusos para melhorar a qualidade da imagem e um algoritmo de otimização, o modelo de aprendizagem desenvolvido superou os principais algoritmos já testados noutros estudos.(18) Fu et al. implementaram também um algoritmo não supervisionado baseado no *Single Shot MultiBox Detector* (SSD), que permite localizar lesões na imagem associada a uma grande base de dados, tornando este estudo o mais representativo da realidade. Os investigadores verificaram que as redes neuronais profundas estudadas foram eficazes na localização de pequenas lesões de carcinoma espinocelular da cavidade oral (CECO) em indivíduos de alto risco, o que pode conduzir a resultados promissores. De acordo com a Organização Mundial de Saúde (OMS), pode ser muito difícil identificar lesões cancerígenas na cavidade oral numa fase inicial, especialmente as que medem menos de dois centímetros e apresentam poucos marcadores visuais. (16) A identificação precisa destas lesões por IA pode ter um impacto terapêutico significativo. Ao detetar lesões com um risco elevado de malignidade o mais cedo possível, os médicos podem tomar decisões mais rápidas e informadas sobre os tratamentos adequados a aplicar a cada doente diretamente no consultório, reduzindo o atraso no tratamento. As fases de pré-processamento da imagem não devem, por conseguinte, ser negligenciadas e devem estar no centro das escolhas efetuadas relativamente ao desenvolvimento da IA, uma vez que são cruciais para a qualidade do seu desempenho. Ferrer et al. e Warin et al. mostraram que a IA pode enfrentar desafios quando se trata de diferenciar estas lesões potencialmente malignas das lesões benignas, com um desempenho inferior ao da deteção de lesões malignas. (3,28)

Isto pode dever-se ao facto de as lesões precoces e as lesões potencialmente malignas poderem partilhar características visuais semelhantes, tornando a distinção mais complexa para os algoritmos de IA. O estudo de Fu et al. também está limitado pela sua base de dados de treino, que não é suficientemente diversificada para permitir a diferenciação entre diferentes lesões. Uma vez que a IA se baseia na aprendizagem a partir de grandes conjuntos de dados anotados, se os dados disponíveis forem limitados ou se as anotações forem menos precisas para as lesões potencialmente malignas, isso pode afetar o desempenho da IA na sua deteção. (16)

Por conseguinte, é crucial continuar a investigação e o desenvolvimento nesta área, a fim de melhorar o desempenho da IA para a deteção de lesões potencialmente malignas na cavidade oral.

McRae et al. tentaram desenvolver uma plataforma que combina IA e amostras de escovas de citologia. Permitindo uma comunicação rápida e concisa entre o laboratório e o dentista através da plataforma, a análise fenotípica da IA mostra bons resultados para a diferenciação de DPMOs, o que a torna uma excelente via para melhorar a precisão do rastreio da IA e ajudar os profissionais de saúde no consultório.(7)

Birur et al., Jubair et al. e Lin et al. tentaram criar um sistema de deteção precoce do cancro adequado a zonas com poucos recursos ou a qualquer profissional de saúde com base na prática, utilizando imagens tiradas com um smartphone.(10,15,31)

Jubair testou o desempenho de um modelo pré-treinado de aprendizagem da luz que poderia potencialmente ser desenvolvido numa aplicação para telemóvel. Esta investigação provou que, apesar da variabilidade da imagem, da potência informática limitada e de um orçamento limitado, era possível desenvolver um algoritmo com boas taxas de previsão para a deteção de CO. (15) Um dos principais desafios da aquisição de imagens de smartphones é a obtenção de imagens com qualidade suficiente para um desempenho ótimo da IA. O desenvolvimento do modelo de IA pode ser seriamente prejudicado pela utilização de fotografias de má qualidade, com impacto na interpretação dos atributos da imagem e conduzindo diretamente a um diagnóstico errado. (21) Lin et al. verificaram que a maioria dos erros de classificação, representando uma percentagem muito pequena, dizia respeito a lesões de baixo e alto risco. Uma vez que estes dois tipos de lesões têm manifestações externas extremamente semelhantes, os resultados são compreensíveis. Por conseguinte, são necessárias imagens de maior qualidade para fazer a distinção. Os investigadores concordam que a IA pode ser um excelente auxiliar de diagnóstico para os médicos de clínica geral, mas não substituirá as competências do especialista. (31)

Birur et al. chegaram à mesma conclusão na sua tentativa de desenvolver um ponto de atendimento especializado para a deteção precoce do cancro oral. Conseguiram uma precisão de 92% ao estabelecerem um protocolo baseado na colaboração entre o pessoal no local, que não é qualificado em patologia oral, e os especialistas. De facto, a equipa no terreno utiliza a IA principalmente para a deteção precoce do cancro e recorre a

especialistas para teleconsulta se a IA for incerta. Sem a intervenção de especialistas, a precisão diminui 5%.(10)

Ao integrar os avanços na deteção da IA nos programas de rastreio, torna-se possível otimizar a utilização dos recursos humanos disponíveis, particularmente em ambientes com poucos recursos, onde o acesso a médicos especialistas pode ser limitado. A integração da inteligência artificial (IA) nestes programas oferece também a vantagem de uma gestão mais eficiente dos recursos. Ao permitir uma identificação mais precisa das pessoas que necessitam de uma avaliação aprofundada, melhora-se a utilização dos recursos médicos disponíveis, concentrando-os nos casos mais críticos. Desta forma, as competências dos profissionais de saúde e as infraestruturas médicas são utilizadas de forma mais eficaz para prestar cuidados específicos onde são mais necessários.

A aplicação da inteligência artificial à deteção precoce da CO não é, por conseguinte, limitada pelo seu custo ou pela complexidade da sua implantação. No entanto, um dos principais desafios que dificultam a sua aplicação e que merece uma atenção especial continua a ser a qualidade e a variabilidade dos dados explorados.

5.3. Desafio da qualidade dos dados e da utilização de variáveis

Os investigadores enfrentam um grande desafio em termos da qualidade dos dados utilizados para treinar modelos de aprendizagem automática no campo da inteligência artificial. Com efeito, para obter um desempenho ótimo, é essencial dispor de conjuntos de dados de elevada qualidade, exaustivos, representativos e corretamente etiquetados. No entanto, pode ser difícil aceder a esses dados, nomeadamente em domínios sensíveis como os cuidados de saúde.

Vários autores procuraram compreender a possível influência da natureza dos dados no desempenho da IA. O tipo de dados utilizados pela IA e a sua associação desempenham um papel fundamental. De facto, Wang et al. descobriram que, ao combinar a análise de imagens intraorais com as informações clínicas do doente, os resultados da análise da IA melhoravam. (26) Da mesma forma, os estudos de Braz et al. e Tseng et al. chegaram a conclusões semelhantes ao combinarem dados da saliva com os registos clínicos dos

doentes. (20,33) Estes resultados sugerem que a aplicação da IA pode ser melhorada através de uma abordagem interdisciplinar, em que diferentes fontes de dados são exploradas em conjunto. Por exemplo, a integração de dados clínicos, fotografias orais e imagiologia poderia permitir à IA obter uma compreensão mais abrangente e exata das lesões e doenças orais. No entanto, é importante sublinhar que são necessários mais estudos para explorar e avaliar a eficácia destas abordagens combinadas.

Tseng et al. procuraram também perceber se o modo de entrada de dados afetava os resultados do modelo de aprendizagem e deduziram que a informação relativa aos anticorpos deveria ser apresentada em formato logarítmico associada à idade em formato normal para obter o melhor desempenho. O trabalho de Tseng et al. realça a importância de ter em conta o modo de entrada dos dados no processo de aprendizagem dos modelos de IA. Ao adaptar o formato das variáveis em função da sua natureza, é possível obter um melhor desempenho e otimizar a utilização da IA. Isto abre mais uma vez o caminho para futuras investigações destinadas a explorar outros aspetos do modo de introdução de dados e a desenvolver abordagens mais sofisticadas para melhorar a eficácia dos modelos de aprendizagem automática.(20)

Ao mesmo tempo, a otimização da utilização das variáveis é também um desafio fundamental. Quando são utilizados conjuntos de dados complexos, pode estar disponível um grande número de variáveis, mas nem todas contribuem significativamente para a tarefa específica que está a ser resolvida. Por conseguinte, é da maior importância identificar as variáveis relevantes e explorá-las de forma otimizada nos modelos de IA, a fim de obter um desempenho de diagnóstico ótimo. Yang et al., depois de analisarem a importância das variáveis detetadas pela IA, identificam que as 5 características mais importantes constituem praticamente toda a informação necessária para a deteção precoce do cancro oral pela IA. (24) O estudo das características da imagem pela IA pode, por conseguinte, ser simplificado e o número de dados reduzido, o que pode reforçar as capacidades dos algoritmos de aprendizagem. Wang et al. fizeram experiências neste domínio e desenvolveram um modelo otimizado que tem em conta apenas as 4 variáveis mais importantes na análise. Esta seleção criteriosa de variáveis resultou num excelente rendimento na deteção precoce. (26) Ao eliminar as variáveis menos significativas, o modelo

é capaz de se concentrar nas informações mais relevantes, o que pode levar a uma maior precisão e eficiência.

Além disso, a fiabilidade dos dados de entrada, em especial as fotografias utilizadas para a IA, desempenha um papel crucial. Dados de entrada de má qualidade, desfocados ou com falhas podem comprometer o desempenho da IA. Lin et al. propuseram-se identificar as características da imagem que influenciam o desempenho dos algoritmos de aprendizagem. Observaram melhores resultados quando a imagem era cortada e centrada para mostrar apenas a região de interesse, ou seja, a lesão, combinada com um protocolo de captura de imagens. (31) Este facto realça a importância de preparar os dados de entrada antes de os utilizar nos modelos de IA. Para responder a este desafio, os profissionais de saúde devem adotar novas medidas. A aquisição de imagens, por exemplo, pode ser normalizada para garantir a coerência das imagens. Isto inclui diretrizes sobre iluminação, ângulo de captura, focagem e posicionamento do doente. Protocolos bem estabelecidos ajudam a minimizar as variações e a obter imagens capazes de melhorar a precisão e a relevância dos resultados obtidos pelos algoritmos de aprendizagem automática. Para tal, é essencial que os dentistas e o pessoal clínico recebam formação sobre a captura de imagens de qualidade para uma utilização otimizada da IA. É necessário que compreendam a importância crucial da precisão da imagem para obter resultados fiáveis. A participação em programas de formação regulares pode ajudar a melhorar as competências e a promover uma prática coerente no campo da inteligência artificial.

As variáveis exploradas pela IA e a qualidade dos dados de entrada não são o único desafio; a composição das amostras representa um verdadeiro desafio para o desenvolvimento da aprendizagem automática. Quando se trata de treinar modelos de IA, é essencial dispor de conjuntos de dados que sejam representativos da população-alvo ou do problema que está a ser estudado. No entanto, a composição das amostras pode frequentemente ser tendenciosa ou não representativa, o que pode levar a um fraco desempenho da IA. Na maioria dos artigos do nosso estudo, existe um desequilíbrio de classes em que algumas classes, como as pessoas com cancro ou as pessoas com DPMO, estão sub-representadas em comparação com outras, como os indivíduos saudáveis, que estão em maioria.

Song et al. tentaram encontrar a melhor abordagem para contrariar este desequilíbrio. A combinação da sobreamostragem das classes minoritárias, da subamostragem das classes maioritárias e da geração de dados sintéticos que permitem aumentar os dados das classes minoritárias gerou uma clara melhoria na redução do desequilíbrio e, por conseguinte, no desempenho do modelo de aprendizagem. (30) É de salientar que a distribuição dos dados utilizados para treinar o algoritmo desempenha um papel fundamental nesta abordagem e merece uma investigação mais aprofundada. Deve ser dada especial atenção à forma como os dados são distribuídos entre as classes minoritárias e maioritárias, de modo a garantir uma representação adequada de cada classe. Este passo é essencial para obter resultados precisos e fiáveis na classificação de imagens de cancro oral.

No entanto, é de notar que a utilização da IA na deteção e classificação do cancro oral apresenta outras limitações e desafios.

5.4. Limites, ética e contexto atual

Todos os artigos mencionados na nossa revisão sistemática são unânimes: a dimensão da base de dados utilizada é demasiado pequena e limita a sua investigação e a correta generalização dos algoritmos. (3,7,10,15–34) Quanto mais pequena for a base de dados, menos representativa é da população e as classes minoritárias não são tidas em conta. Uma das principais razões para a má classificação do modelo de treino é o facto de os dados de treino não representarem a diversidade dos casos reais devido à sua pequena dimensão. Se as imagens utilizadas para a formação não abrangerem uma gama suficientemente ampla de lesões ou condições específicas, a IA pode ter dificuldade em reconhecer e classificar corretamente estes casos. De facto, Jerayaj demonstrou que, ao aumentar o número de imagens na sua base de dados, em particular as imagens de cancro, a precisão do modelo aumentou 5%. (32) Além disso, de acordo com Song et al, as medidas de precisão não são adequadas para analisar o desempenho da IA para classes minoritárias ou classes com desequilíbrios significativos, como as classes com lesões de alto risco. (30) Por conseguinte, é importante melhorar a representatividade dos dados de treino, incluindo uma vasta gama de situações clínicas e possíveis variações. Isto pode implicar a recolha de dados adicionais de diferentes fontes, a colaboração com outras instituições de saúde ou a

criação de bases de dados mais completas e diversificadas. Por conseguinte, é necessário efetuar estudos com uma maior dimensão da amostra e analisar as medidas mais adequadas para quantificar o desempenho da IA.

Outra limitação importante é a falta de validação externa dos estudos. A maioria dos estudos apenas efetuou uma validação interna e apenas 6 incluíram uma validação externa, através da inclusão de dados de vários hospitais diferentes. (10,15,16,21,30) A recolha de dados adicionais de diferentes fontes e de outras clínicas, como a validação externa, pode ajudar a diversificar a amostra de validação. Isto pode permitir a obtenção de dados de uma variedade de populações com diferentes características demográficas, clínicas ou geográficas. Essa diversidade pode, por conseguinte, garantir uma avaliação mais sólida do desempenho da IA.

Para além da combinação de dados clínicos e imagens, nenhum dos artigos da nossa revisão estudou uma abordagem multidisciplinar, combinando dados de diferentes tipos, como a combinação de imagens e fotografias orais ou o trabalho de vários profissionais de saúde qualificados numa área diferente. A integração de diferentes fontes de dados, como imagens médicas, fotografias orais, dados clínicos, informações histopatológicas e biomarcadores, pode fornecer uma visão mais abrangente e aprofundada da doença. Esta abordagem multidisciplinar permitiria explorar todo o potencial da IA, combinando os pontos fortes de cada modalidade de dados para uma maior precisão do diagnóstico.

Para implementar uma abordagem multidisciplinar, é essencial estabelecer uma colaboração entre vários profissionais de saúde, incluindo dentistas, radiologistas, patologistas e investigadores especializados em cancro oral. Esta colaboração favoreceria o intercâmbio de conhecimentos, a partilha de dados e o desenvolvimento de protocolos de trabalho harmonizados e comuns.

De facto, os investigadores Heo et al. e Wang et al. adotaram uma abordagem normalizada para a aquisição de dados nos seus respetivos estudos. (21,26) Reconheceram a importância de implementar protocolos coerentes e reprodutíveis para garantir a fiabilidade dos resultados. Graças a esta normalização, observaram uma melhoria no desempenho dos seus modelos de IA e puderam avaliar os seus resultados de forma mais objetiva. Esta abordagem

ajuda a reforçar a credibilidade dos estudos e facilita a comparação de diferentes tipos de investigação no domínio da deteção precoce do cancro oral.

Para aumentar as bases de dados e facilitar a colaboração multidisciplinar, convém considerar a introdução de um protocolo obrigatório para as fotografias orais, semelhante ao que existe para as radiografias orais durante os procedimentos dentários. Este protocolo padronizado poderia ser incorporado na prática clínica de rotina, exigindo que as fotografias fossem tiradas sistematicamente durante os exames orais regulares ou durante procedimentos específicos. Wang et al. tentaram facilitar a utilização generalizada da IA, desenvolvendo um sítio Web acessível através do sítio OPMDRisk (<http://www.opmd-risk.com>) para criar um apoio personalizado ao doente, acessível a todos os profissionais de saúde. Ainda se aguardam os resultados finais deste estudo, mas ele representa um passo importante no sentido de vulgarizar a utilização da IA na saúde oral. (26)

A integração de um protocolo deste tipo permitiria a criação de uma vasta base de dados de imagens orais, fornecendo informações visuais exaustivas sobre a cavidade oral e os tecidos circundantes. Estas imagens poderiam ser anonimizadas, armazenadas de forma segura e disponibilizadas aos profissionais de saúde para posterior análise e estudos de investigação.

Antes da criação de um sistema deste tipo, há outros limites à aplicação da IA na deteção precoce do cancro oral que devem ser tidos em conta.

Na maioria dos artigos estudados na revisão relativos à IA supervisionada, os investigadores utilizaram peritos humanos para rotular os dados de treino. No entanto, é importante reconhecer que mesmo os peritos humanos podem apresentar alguma variação na interpretação e anotação dos dados. Fatores como a subjetividade, a experiência individual e os preconceitos profissionais podem influenciar as anotações feitas pelos investigadores. No artigo de Yuan et al., o nível de experiência dos peritos nem sequer é mencionado. (17) Por conseguinte, é necessário adotar medidas adequadas para garantir a qualidade e a fiabilidade das anotações.

De um ponto de vista ético, a utilização da inteligência artificial está também a causar grande controvérsia com a noção de " *black box* ". Os algoritmos de IA podem tomar decisões complexas sem fornecer uma explicação clara do seu processo de decisão, que

muitas vezes escapa à compreensão humana. Esta opacidade pode ser motivo de preocupação no sector da saúde, onde as decisões tomadas pela IA podem ter consequências diretas na vida das pessoas. De acordo com Wang F. et al, a falta de transparência suscita apreensões quanto à confiança e fiabilidade dos sistemas de IA. Os profissionais de saúde, os investigadores e os legisladores são confrontados com o desafio de conciliar a utilização destas tecnologias para melhorar os cuidados de saúde com a necessidade de compreender e justificar as decisões tomadas por estes sistemas. (36)

A isto acresce a noção da responsabilidade da IA pelas decisões terapêuticas tomadas. Como médicos dentistas, é crucial clarificar as responsabilidades associadas à utilização da inteligência artificial (IA) na nossa prática. De acordo com Naik et al, os sistemas de IA podem ser utilizados para apoiar o diagnóstico e o tratamento dos doentes, mas em caso de danos ou erros, é essencial determinar quem é responsável. Para tal, é necessário clarificar os papéis e as responsabilidades dos criadores de algoritmos, dos fornecedores de tecnologia, dos profissionais de saúde e das entidades reguladoras. Além disso, os dentistas devem ser capazes de explicar aos doentes como a IA é utilizada no seu tratamento e informá-los sobre as limitações e os potenciais riscos associados. (37)

Outra consideração ética diz respeito à confidencialidade e à proteção dos dados. A utilização da IA para a deteção precoce do cancro oral exige a recolha e o armazenamento de grandes quantidades de dados médicos sensíveis com o consentimento informado do doente. Os doentes devem receber informações claras e pormenorizadas sobre a finalidade para a qual os seus dados serão recolhidos, os tipos de dados que serão recolhidos e as medidas adotadas para garantir a confidencialidade das suas informações pessoais. Os profissionais de saúde têm a obrigação de garantir a segurança dos dados dos seus doentes e de respeitar as regras de confidencialidade. É essencial estar consciente dos riscos potenciais associados à divulgação não autorizada dos registos dos doentes. De acordo com a investigação de Naik et al., a utilização da IA envolve frequentemente a transferência e o armazenamento de dados em servidores externos, o que pode aumentar o risco de violação da privacidade. (37)

Os modelos de aprendizagem devem, por conseguinte, ser avaliados e validados de acordo com as recomendações da Associação para o Avanço da Inteligência Artificial, tal como

salientado no artigo de Naik et al.(37) É essencial criar sistemas de acompanhamento e controlo para garantir que a utilização da IA respeita os princípios éticos e as normas de proteção de dados, bem como para medir e avaliar a sua fiabilidade e desempenho de forma lógica e estatística. Devem também ser efetuadas avaliações regulares para medir o impacto da IA na privacidade e nos direitos fundamentais dos indivíduos. (37)

Por último, é importante sublinhar que a responsabilidade não deve recair apenas sobre os algoritmos, mas também sobre os profissionais de saúde que os utilizam. Estes devem estar conscientes das suas próprias responsabilidades e não procurar escapar-lhes culpando apenas a IA. (37)

Os avanços no âmbito da inteligência artificial estão em constante evolução, e o quadro jurídico para a sua utilização ainda não está totalmente estabelecido. Consciente desta situação, a Comissão Europeia emitiu em 2021 uma proposta, que se encontra atualmente em estudo, para regulamentar a utilização da inteligência artificial na União Europeia. Esta proposta visa estabelecer normas e regras para enquadrar a utilização da IA, com particular atenção à salvaguarda dos direitos fundamentais, à promoção da transparência, à definição de responsabilidades e à garantia da segurança dos sistemas de IA. (38)

Para além das limitações decorrentes da utilização da inteligência artificial, esta revisão sistemática também tem as suas próprias limitações. Em primeiro lugar, enfrenta os desafios inerentes às revisões, como a heterogeneidade dos estudos incluídos, a qualidade dos estudos e a complexidade dos dados analisados. Para além disso, foram identificadas considerações específicas a esta revisão. Por exemplo, os dados utilizados provêm de estudos que utilizaram diferentes técnicas de diagnóstico, o que levou a alguma variação nos resultados e à necessidade de avaliar a sensibilidade de cada método de deteção de forma independente. Além disso, devem ser tidas em conta potenciais diferenças na qualidade do equipamento de aquisição de dados utilizado nos diferentes estudos, bem como variações nos protocolos e procedimentos, que podem influenciar a precisão do diagnóstico. A inclusão de diferentes tipos de cancro com diferentes localizações no estudo pode introduzir níveis heterogéneos de dificuldade de deteção, o que pode potencialmente afetar os resultados. É essencial ter em conta esta diversidade ao avaliar o desempenho da

IA, a fim de evitar vieses e obter uma compreensão mais exata da sua eficácia na deteção de diferentes tipos de cancro.

O campo da IA está a evoluir rapidamente, com novos artigos a serem publicados em permanência. Isto significa que algumas informações podem tornar-se rapidamente obsoletas, uma vez que estão constantemente a ser feitas novas descobertas. Por conseguinte, é importante manter-se a par dos últimos desenvolvimentos, consultando fontes de informação atualizadas e fiáveis.

6. Conclusões

Depois de analisar cuidadosamente os vários estudos e investigações sobre a aplicação da inteligência artificial (IA) no diagnóstico precoce do cancro oral, esta abordagem de diagnóstico oferece perspectivas promissoras para melhorar a precisão, a rapidez e a eficiência do rastreio, ajudando assim a preservar e a melhorar o prognóstico de saúde dos doentes.

No entanto, é importante reconhecer que a utilização da IA também levanta desafios éticos e práticos. Entre estes desafios, a questão da ética e da responsabilidade na utilização da IA, bem como a necessidade de garantir que as decisões tomadas pelos sistemas de IA sejam transparentes e compreensíveis, são pontos cruciais a considerar.

Além disso, para garantir uma utilização ótima da inteligência artificial (IA) na prática clínica, é essencial integrar colaboração das diferentes modalidades de aquisição de dados, como a imagiologia, dados de TCO, fotografia oral e biomarcadores permitindo criar perfis completos dos doentes que possam ser utilizados por todos os profissionais de saúde. Isto requer uma abordagem multidisciplinar e a harmonização de protocolos, a fim de explorar plenamente o potencial da IA na avaliação diagnóstica e na deteção precoce do cancro oral.

Apesar destes desafios, é importante notar que a IA continua a evoluir rapidamente, oferecendo novas oportunidades e abrindo perspectivas para o diagnóstico precoce do cancro oral. É essencial permanecer atento aos últimos avanços científicos e realizar estudos rigorosos para avaliar a eficácia, a segurança e as implicações a longo prazo da utilização da IA neste domínio. Por exemplo, ao desenvolver a identificação de novos biomarcadores utilizando modelos de aprendizagem, o que poderia contribuir para o desenvolvimento de novas terapêuticas direcionadas e de medicamentos mais eficazes. Ao compreenderem melhor os mecanismos biológicos subjacentes ao cancro, os investigadores poderão conceber terapêuticas mais específicas que atuem diretamente sobre esses biomarcadores, melhorando assim os resultados do tratamento e, conseqüentemente, a sobrevivência dos doentes, conduzindo o sector da saúde a uma nova medicina personalizada.

7. Referencias bibliográficas

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* Maio de 2021;71(3):209–49.
2. Güneri P, Epstein JB. Why are we still unable to accurately determine the malignant potential or the behavior of oral mucosal lesions? *Oral Oncol.* Agosto de 2017;71:177–9.
3. Warin K, Limprasert W, Suebnukarn S, Jinaporntham S, Jantana P, Vicharueang S. AI-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer. *PLoS One.* Agosto de 2022;17.
4. Bagan J, Sarrion G, Jimenez Y. Oral cancer: clinical features. *Oral Oncol.* Junho de 2010;46(6):414–7
5. Seoane J, Takkouche B, Varela-Centelles P, Tomás I, Seoane-Romero JM. Impact of delay in diagnosis on survival to head and neck carcinomas: a systematic review with meta-analysis. *Clin Otolaryngol.* Abril de 2012;37(2):99–106.
6. Warnakulasuriya S, Kujan O, Aguirre-Urizar JM, Bagan J V., González-Moles MÁ, Kerr AR, et al. Oral potentially malignant disorders: A consensus report from an international seminar on nomenclature and classification, convened by the WHO Collaborating Centre for Oral Cancer. *Oral Dis.* Novembro de 2021;27(8):1862–80.
7. McRae MP, Modak SS, Simmons GW, Trochesset DA, Kerr AR, Thornhill MH, et al. Point-of-care oral cytology tool for the screening and assessment of potentially malignant oral lesions. *Cancer Cytopathol.* Março de 2020;128(3):207–20.

8. Amisha, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. *J Family Med Prim Care*. 2019;8(7):2328.
9. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med*. Abril de 2019;380(14):1347–58.
10. Birur N P, Song B, Sunny SP, G K, Mendonca P, Mukhia N, et al. Field validation of deep learning based Point-of-Care device for early detection of oral malignant and potentially malignant disorders. *Sci Rep*. Dezembro de 2022;12(1).
11. García-Pola M, Pons-Fuster E, Suárez-Fernández C, Seoane-Romero J, Romero-Méndez A, López-Jornet P. Role of artificial intelligence in the early diagnosis of oral cancer. A scoping review. *Cancers (Basel)*. Setembro de 2021;13(18):4600–24.
12. Ilhan B, Guneri P, Wilder-Smith P. The contribution of artificial intelligence to reducing the diagnostic delay in oral cancer. *Oral Oncol*. Maio de 2021;116:105254.
13. Ilhan B, Lin K, Guneri P, Wilder-Smith P. Improving Oral Cancer Outcomes with Imaging and Artificial Intelligence. *J Dent Res*. Março de 2020;99(3):241.
14. Araújo ALD, da Silva VM, Kudo MS, de Souza ESC, Saldivia-Siracusa C, Giraldo-Roldán D, et al. Machine learning concepts applied to oral pathology and oral medicine: A convolutional neural networks' approach. *Journal of Oral Pathology and Medicine*. Janeiro de 2023;
15. Jubair F, Al-karadsheh O, Malamos D, Al Mahdi S, Saad Y, Hassona Y. A novel lightweight deep convolutional neural network for early detection of oral cancer. *Oral Dis*. Março de 2021;28(4):1123–30.
16. Fu Q, Chen Y, Li Z, Jing Q, Hu C, Liu H, et al. A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study. *EClinicalMedicine*. Outubro de 2020;27:100558.

17. Yuan W, Yang J, Yin B, Fan X, Yang J, Sun H, et al. Noninvasive diagnosis of oral squamous cell carcinoma by multi-level deep residual learning on optical coherence tomography images. *Oral Dis.* Janeiro de 2022;
18. Marzouk R, Alabdulkreem E, Dhahbi S, Nour MK, Duhayyim M Al, Othman M, et al. Deep Transfer Learning Driven Oral Cancer Detection and Classification Model. *Computers, Materials and Continua.* Janeiro de 2022;73(2):3905–20.
19. Hsieh CH, Chen WM, Hsieh YS, Fan YC, Yang PE, Kang ST, et al. A novel multi-gene detection platform for the analysis of miRNA Expression. *Sci Rep.* Dezembro de 2018;8(1).
20. Tseng YJ, Wang YC, Hsueh PC, Wu CC. Development and validation of machine learning-based risk prediction models of oral squamous cell carcinoma using salivary autoantibody biomarkers. *BMC Oral Health.* Dezembro de 2022;22(1).
21. Heo J, Lim JH, Lee HR, Jang JY, Shin YS, Kim D, et al. Deep learning model for tongue cancer diagnosis using endoscopic images. *Sci Rep.* Dezembro de 2022;12(1).
22. Aubreville M, Knipfer C, Oetter N, Jaremenko C, Rodner E, Denzler J, et al. Automatic Classification of Cancerous Tissue in Laserendomicroscopy Images of the Oral Cavity using Deep Learning. *Sci Rep.* Dezembro de 2017;7(1).
23. Shao S, Mao N, Liu W, Cui J, Xue X, Cheng J, et al. Epithelial salivary gland tumors: Utility of radiomics analysis based on diffusion-weighted imaging for differentiation of benign from malignant tumors. *J Xray Sci Technol.* 2020;28(4):799–808.
24. Yang Z, Shang J, Liu C, Zhang J, Liang Y. Classification of oral salivary gland tumors based on texture features in optical coherence tomography images. *Lasers Med Sci.* Março de 2022;37(2):1139–46.

25. Gunduz E, Alçin OF, Kizilay A, Yildirim IO. Deep learning model developed by multiparametric MRI in differential diagnosis of parotid gland tumors. *European Archives of Oto-Rhino-Laryngology*. Novembro de 2022;279(11):5389–99.
26. Wang X, Yang J, Wei C, Zhou G, Wu L, Gao Q, et al. A personalized computational model predicts cancer risk level of oral potentially malignant disorders and its web application for promotion of non-invasive screening. *J Oral Pathol Med*. Maio de 2020;49(5):417–26.
27. He Z, Mao Y, Lu S, Tan L, Xiao J, Tan P, et al. Machine learning-based radiomics for histological classification of parotid tumors using morphological MRI: a comparative study. *Eur Radiol*. Dezembro de 2022;32(12).
28. Ferrer-Sánchez A, Bagan J, Vila-Francés J, Magdalena-Benedito R, Bagan-Debon L. Prediction of the risk of cancer and the grade of dysplasia in leukoplakia lesions using deep learning. *Oral Oncol*. Setembro de 2022;132.
29. Mentel S, Gallo K, Wagendorf O, Preissner R, Nahles S, Heiland M, et al. Prediction of oral squamous cell carcinoma based on machine learning of breath samples: a prospective controlled study. *BMC Oral Health*. Dezembro de 2021;21(1).
30. Song B, Li S, Sunny S, Gurushanth K, Mendonca P, Mukhia N, et al. Classification of imbalanced oral cancer image data from high-risk population. *J Biomed Opt*. Outubro de 2021;26(10).
31. Lin H, Chen H, Weng L, Shao J, Lin J. Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis. *J Biomed Opt*. Agosto de 2021;26(8).
32. Jeyaraj PR, Samuel Nadar ER. Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm. *J Cancer Res Clin Oncol*. Abril de 2019;145(4):829–37.

33. Braz DC, Neto MP, Shimizu FM, Sá AC, Lima RS, Gobbi AL, et al. Using machine learning and an electronic tongue for discriminating saliva samples from oral cavity cancer patients and healthy individuals. *Talanta*. Junho de 2022;243.
34. Alhazmi A, Alhazmi Y, Makrami A, Masmali A, Salawi N, Masmali K, et al. Application of artificial intelligence and machine learning for prediction of oral cancer risk. *J Oral Pathol Med*. Maio de 2021;50(5):444–50.
35. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. 2009;
36. Wang F, Kaushal R, Khullar D. Should Health Care Demand Interpretable Artificial Intelligence or Accept «Black Box» Medicine? *Ann Intern Med*. Janeiro de 2020;172(1):59–61.
37. Naik N, Hameed BMZ, Shetty DK, Swain D, Shah M, Paul R, et al. Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility? *Front Surg*. Março de 2022;9:862322.
38. PT COMISSÃO EUROPEIA REGULAMENTO DO PARLAMENTO EUROPEU E DO CONSELHO QUE ESTABELECE REGRAS HARMONIZADAS EM MATÉRIA DE INTELIGÊNCIA ARTIFICIAL (REGULAMENTO INTELIGÊNCIA ARTIFICIAL) E ALTERA DETERMINADOS ATOS LEGISLATIVOS DA UNIÃO.