República Federativa do Brasil

Ministério do Desenvolvimento, Indústria, Comércio e Serviços

Instituto Nacional da Propriedade Industrial

* B R 1 0 2 0 2 5 0 0 2 6 3 7 A 2 *

(21) BR 102025002637-6 A2

(22) Data do Depósito: 10/02/2025

(43) Data da Publicação Nacional: 18/08/2026

(54) Título: MÉTODOS E SISTEMA PARA DIAGNÓSTICO DE DISFAGIA OROFARÍNGEA BASEADO EM INTELIGÊNCIA ARTIFICIAL

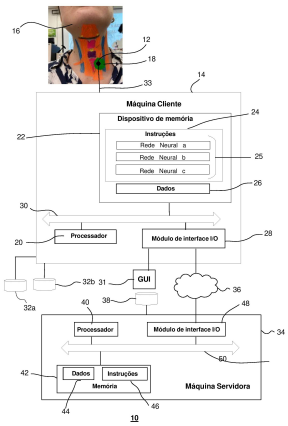
(51) Int. Cl.: A61B 8/00; A61B 8/12.

(52) CPC: A61B 8/00; A61B 8/52; A61B 8/5207; A61B 8/12.

(71) Depositante(es): FUNDACAO UNIVERSIDADE FEDERAL DE CIENCIAS DA SAUDE DE PORTO ALEGRE; UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS).

(72) Inventor(es): RAFAELA SOARES RECH; FERNANDO NEVES HUGO; CARLA DINIZ LOPES BECKER; TIAGO BECKER; DIEINE ESTELA BERNIERI SCHIAVON DE ANDRADE.

(57) Resumo: MÉTODOS E SISTEMA PARA DIAGNÓSTICO DE DISFAGIA OROFARÍNGEA BASEADO EM INTELIGÊNCIA ARTIFICIAL. Um sistema de suporte para o diagnóstico de disfagia orofaríngea, o sistema compreendendo um ou mais processadores acoplados a pelo menos um dispositivo de memória armazenando instruções executáveis por um ou mais processadores para pelo menos executar as operações de: com um dispositivo sensor, adquirir sinais acústicos emitidos durante um evento de deglutição de um material por uma pluralidade de indivíduos; identificar sinais acústicos de deglutição dentro dos sinais acústicos emitidos; préprocessar os sinais acústicos de deglutição; converter os sinais acústicos de deglutição em dados de imagem; extrair aspectos representativos do evento de deglutição dos dados de imagem; gerar um conjunto de dados compreendendo os dados de imagem; dividir os dados de imagem em um conjunto de dados de treinamento e um conjunto de dados de teste; inserir o conjunto de dados de treinamento em uma rede neural para treinar um modelo para classificar os dados de imagem com base em pelo menos um dos aspectos, como indicativo de deglutição normal e deglutição anormal devido à disfagia orofaríngea; validar o diagnóstico de disfagia orofaríngea usando dados de teste; e produzir um modelo treinado para pelo menos diagnosticar disfagia (...).



MÉTODOS E SISTEMA PARA DIAGNÓSTICO DE DISFAGIA OROFARÍNGEA BASEADOS EM INTELIGÊNCIA ARTIFICIAL

CAMPO DA INVENÇÃO

[0001] A presente invenção pertence ao campo do uso de inteligência artificial na medicina e refere-se a métodos e sistema para diagnóstico de disfagia orofaríngea.

DESCRIÇÃO DO ESTADO DA TÉCNICA

[0002] A função da deglutição desempenha um papel vital na vida diária de todos os indivíduos, transportando nutrientes e energia necessária da cavidade oral para o estômago, garantindo que nenhuma substância entre nas vias aéreas e chegue aos pulmões. A deglutição saudável requer um bom controle muscular das estruturas da cavidade oral, faringe, laringe e esôfago, para que o alimento ingerido seja reduzido a uma consistência adequada que possa ser transportada da boca para o estômago com segurança e sem desconforto. Para alcançar este objetivo, todas as estruturas mecânicas e neurológicas precisam estar intactas e saudáveis para que o processo ocorra adequadamente.

[0003] Os distúrbios da deglutição são caracterizados por problemas que envolvem a cavidade oral, faringe, esôfago ou junção gastroesofágica. Tais distúrbios são classificados como qualquer dificuldade no processo de deglutição, começando com a formação do bolo alimentar na boca e sua

subsequente propulsão pela faringe, esôfago ou estômago. As alterações anatômicas e fisiológicas são classificadas como condição digestiva na Classificação Internacional de Doenças (CID), apresentadas como CID-9 (787.2) e CID-10 (R13) e comumente referidas como disfagia.

[0004] Os sintomas que acompanham a disfagia incluem tosse ou engasgo com alimentos ou saliva, pneumonia recorrente, refluxo gastroesofágico, febre sem causa aparente, sensação de comida presa na garganta, recusa de comida e sonolência durante as refeições. Os sinais clínicos característicos da aspiração podem ser ausência de tosse, voz com qualidade vocal úmida, dispneia ou aumento de secreção nas vias aéreas superiores. Além disso, outros sinais clínicos esperados para um indivíduo disfágico são dificuldade em agarrar os alimentos por via oral, dificuldade em mastigar, controle oral reduzido de alimentos na cavidade oral, escape de alimentos para a faringe, regurgitação nasal ou oral, lentidão no manuseio dos alimentos e preparação dos alimentos para deglutição, engasgos frequentes, insegurança e ansiedade nas refeições, qualidade vocal úmida, hipersecreção constante e alterações no padrão respiratório.

[0005] A disfagia é considerada um problema de saúde pública, uma síndrome geriátrica, com uma prevalência estimada de 45,3% em pessoas idosas e aproximadamente 20% da população em geral. Suas consequências resultam em alta morbidade e mortalidade. Além disso, essa condição resulta em aumento de custos para os sistemas de saúde. Essa

condição representa um importante indicador de saúde para a população idosa, pois além de ser um dos sintomas de várias doenças prevalentes nesse segmento populacional, também pode estar associada à morbidade, mortalidade precoce e pode levar a diversas complicações clínicas, incluindo desidratação, desnutrição e pneumonia aspirativa. Além disso, o envelhecimento não saudável pode estar associado à sua presença. As alterações na deglutição são progressivas com a idade, com aumento da fragilidade e diminuição da independência. A disfagia muitas vezes não é reconhecida e, portanto, subdiagnosticada em diferentes estabelecimentos de saúde desse segmento populacional.

[0006] A disfagia é um sintoma de várias doenças prevalentes na população (doenças crônicas e neurológicas). As principais doenças relacionadas à disfagia são demência, doença de Alzheimer, doença de Parkinson, acidente vascular cerebral (AVC), câncer, doença pulmonar obstrutiva crônica, esclerose múltipla, esclerose lateral amiotrófica, divertículo de Zenker, distrofia muscular, lesão cerebral traumática, entre outras.

[0007] Para minimizar os efeitos adversos da disfagia na saúde dos indivíduos, a detecção precoce é essencial. Em pacientes hospitalizados, a identificação rápida de pacientes disfágicos está associada à redução do risco de pneumonia, à redução do tempo de internação hospitalar e à melhor relação custo-efetividade associada à redução das internações hospitalares. Em pacientes ambulatoriais, a

rápida identificação pode prevenir a hospitalização e o agravamento do problema de saúde.

[0008] Há uma tentativa de instituições que representam a fonoaudiologia e áreas afins de padronizar e validar o conjunto de avaliações clínicas fonoaudiológicas, que avançaram substancialmente nos últimos trinta anos, à medida que a maioria das instituições e profissionais adota diferentes métodos, instrumentos e consistências alimentares de acordo com suas necessidades e preferências.

[0009] Os métodos diagnósticos existentes incluem avaliação clínica, avaliação videofluoroscópica e avaliação endoscópica da deglutição, que são realizadas com testes alimentares e visam analisar a eficiência e a segurança dos indivíduos avaliados. Além desses métodos estão a manometria faringoesofágica, eletromiografia de superfície, ultrassom, cintilografia, sonar doppler, endoscopia virtual por tomografia computadorizada com software de reconstrução e biossusceptometria de corrente alternada (BAC).

[0010] Geralmente, a videofluoroscopia da deglutição é considerada o exame de referência para o diagnóstico de disfagia. Também é conhecido como exame de deglutição de bário modificado e captura imagens vídeo-radiográficas sequenciais de alimentos e líquidos misturados com contraste de bário à medida que são transportados através da cavidade oral, cavidade faríngea e esôfago em tempo real. No entanto, o exame deve ser realizado em instalações designadas por pessoal treinado que esteja sujeito à exposição à radiação.

Além disso, a avaliação é mais qualitativa e requer conhecimento especializado para interpretá-la, e a videofluoroscopia da deglutição é substancialmente cara e pode não estar disponível na maioria dos serviços de saúde. Além disso, a solução de bário é conhecida por ser prejudicial ao meio ambiente.

[0011] Na prática cotidiana, a avaliação clínica da deglutição realizada por um fonoaudiólogo permanece essencial, destacando-se como a abordagem mais viável por sua praticidade, baixo custo e pela especialização do profissional. Esse procedimento abrange a análise de condições anatômicas e fisiológicas (orofaciais, neurológicas e corporais), além da observação de sinais e sintomas relacionados à deglutição em todas as consistências alimentares possíveis. A avaliação oferece menos riscos aos pacientes, por exemplo, por não os expor aos raios-X. No entanto, essa avaliação é frequentemente contestada, pois carece de indicadores quantitativos objetivos. Embora a avaliação clínica, essencial na prática, evite exames complementares, os métodos atuais têm uma acurácia muito baixa, o que torna a videofluoroscopia desejável.

[0012] Entre as formas de diagnóstico, a análise acústica revelou o potencial de aumentar a acurácia diagnóstica e qualificar a avaliação clínica. Os sons de deglutição são compostos por três componentes, a saber:

um primeiro sinal fraco, que corresponde a elevação laríngea e anteriorização do osso hióide e a passagem do

bolo alimentar através da orofaringe e/ou hipofaringe;

um segundo sinal forte, que está relacionado com a abertura da transição faringoesofágica; e,

um terceiro ruído, um sinal fraco que está associado à laringe, resultante da sua descida e abertura, bem como a passagem do bolo alimentar através da transição faringoesofágica.

[0013] No entanto, ainda não há consenso na literatura quanto à correspondência fisiológica dos sons da deglutição. A percepção humana desses sons pode ser muito difícil de perceber, considerando as características clínicas de cada caso, o ambiente, a qualidade do equipamento e a experiência do examinador.

Panorama tecnológico

[0014] A ausculta cervical tem sido usada como uma parte importante do diagnóstico clínico; no entanto, mesmo os sons obtidos por estetoscópios digitais estão sujeitos à subjetividade da decisão clínica do avaliador. O uso de parâmetros para modificação da ausculta cervical durante e após a deglutição é cuidadosamente analisado em protocolos clínicos para evitar viés de aferição. Além disso, há necessidade de padronização dos métodos de investigação dos sons da deglutição para análise adequada da acurácia dos parâmetros investigados.

[0015] Os parâmetros acústicos da deglutição podem contribuir para uma análise mais precisa do diagnóstico

clínico da disfagia, possibilitando a identificação das principais características das ondas captadas pelos sons da deglutição e reduzindo a subjetividade da avaliação clínica realizada pelo fonoaudiólogo e outros profissionais de saúde.

[0016] Para a análise acústica da deglutição, vários métodos podem ser usados como detectores acústicos. Os instrumentos de amplificação podem ser usados para capturar som fornecer análise acústica do som através de gráficos que fornecem visualmente a amplitude, duração e frequência do som de deglutição. O estetoscópio digital, por exemplo, é uma alternativa que permite armazenar informações sonoras sobre a deglutição em um banco de dados.

[0017] A ausculta cervical é usada na avaliação clínica para avaliar a fase faríngea da deglutição através da percepção dos sons da deglutição usando um instrumento de amplificação de som. O instrumento serve para determinar a segurança da deglutição, mas também oferece informações como o tempo desses sons que foram detectados e amplificados. A ausculta cervical digital reduz a subjetividade do método tradicional ao apresentar uma descrição dos sons encontrados durante os exames e o tempo de trânsito do bolo alimentar, juntamente com a amplitude e frequência da onda encontrada na análise. Além disso, a ausculta cervical digital é um método de avaliação clínica de custo relativamente baixo.

[0018] Técnicas mais econômicas, não invasivas, potencialmente precisas e fáceis de coletar foram estudadas

nos últimos anos para auxiliar no diagnóstico de várias patologias (Alzheimer, Parkinson, Covid-19, Disfagia, entre outras). Essas técnicas contam com o uso de inteligência artificial, mais especificamente, deep learning para a tarefa de classificação com base em características extraídas de sinais acústicos (fala, tosse, deglutição).

[0019] Em 2014, Jayatilake e colaboradores desenvolveram um produto chamado “Swallowscope” e propuseram um algoritmo para detectar a atividade de deglutição em 15 indivíduos com base na deglutição de saliva. Os autores concluíram que este dispositivo, baseado em inteligência artificial utilizada a partir de telefones móveis, poderia ser uma importante ferramenta quantitativa para maior acurácia diagnóstica da disfagia.

[0020] Em 2015, os pesquisadores testaram o produto Swallowscope em 15 indivíduos saudáveis e 70 com diagnóstico prévio de disfagia, demonstrando sua capacidade de reconhecer com precisão as características temporais das atividades de deglutição. Essa pesquisa deu origem à patente JP2019051129A.

[0021] Esse mesmo grupo de pesquisadores aprimorou sua técnica e, em 2020, propôs um dispositivo de monitoramento de deglutição em tempo real baseado em smartphone, o Gokuri, da PLIMES, Inc., Japão (WO2022085188). O Gokuri é um dispositivo portátil e integrável ao corpo que se conecta a um telefone móvel (smartphone) para fornecer uma avaliação em tempo real das funções de deglutição usando sons de

deglutição. A avaliação compreende o monitoramento do movimento da deglutição, sinais vitais, eventos como tosse, engasgos, tempo de deglutição e sons da deglutição pelos profissionais de saúde. O produto consiste em três unidades principais:

1) uma cinta de pescoço vestível com um microfone que captura sons de deglutição,

2) um amplificador, e

3) uma interface de smartphone que mostra resultados e estatísticas.

[0022] Em 2022, Subramani e seus colaboradores apresentaram um método baseado em uma rede neural convolucional profunda para segmentação de pixels semânticos para classificar sinais de deglutição saudáveis e disfágicos, através do mel-espectrograma. Os sons de deglutição foram gravados de um total de 24 indivíduos usando um sistema de ausculta cervical. O modelo de desempenho dos autores alcançou um F-score médio de 80,13%.

[0023] A pesquisa de Kim e colaboradores, também em 2022, aplicou técnicas de deep learning para classificar a disfagia através da análise de sinais gerados em videofluoroscopia de 190 participantes. Uma rede neural convolucional foi usada para classificar os achados videofluoroscópicos (deglutição normal, penetração laríngea e aspiração). A classificação foi determinada considerando picos (altos e baixos). Os resultados para o conjunto de

dados utilizados na avaliação foram uma área sob a curva (Curva ROC) de 0,942 para achados normais, 0,878 para penetração laríngea e 1,0 para aspiração. Os autores concluíram que a rede neural convolucional pode ser aplicada para detectar a presença de penetração e aspiração laríngea por meio de sinais coletados de exames de videofluoroscopia de pacientes com disfagia.

Referência não-patentária

[0024] Jayatilake, D., Suzuki, K., Teramoto, Y., Ueno, T., Nakai, K., Hidaka, K., Ayuzawa ,S., Eguchi, K., & Matsumura, A. (2014). Swallowscope: A smartphone based device for the assessment of swallowing ability. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)* (pp. 697-700). IEEE.

Jayatilake, D., Ueno, T., Teramoto, Y., Nakai, K., Hidaka, K., Ayuzawa, S., Eguchi, K., Matsumura, A., & Suzuki, K. (2015). Smartphone-Based Real-time Assessment of Swallowing Ability From the Swallowing Sound. In *IEEE J Transl Eng Health Med.* ;3:2900310. doi: 10.1109/JTEHM.2015.2500562. PMID: 27170905; PMCID: PMC4848072.

Subramani, S., Rao, A., Roy, A., Hegde, P. S., & Ghosh, P. K. (2022). Segnet-based deep representation learning for dysphagia classification. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1141-1145). IEEE.

Kim, J. K., Choo, Y. J., Choi, G. S., Shin, H., Chang, M. C., & Park, D. (2022). Deep learning analysis to automatically detect the presence of penetration or aspiration in videofluoroscopic swallowing study. *Journal of Korean Medical Science*, *37*(6).

Problema a ser resolvido

[0025] Existe um valor substancial no desenvolvimento de uma ferramenta facilmente implementável e escalável para o diagnóstico e vigilância da disfagia. Essa abordagem é valiosa em países de baixa e média renda, onde a capacidade de detecção de casos é limitada, onde a captura de dados de sintomas por meio dos serviços de saúde é insuficiente devido à escassez de profissionais especializados na área e onde os métodos diagnósticos são muito heterogêneos.

[0026] Há também um valor substancial no desenvolvimento de uma ferramenta para os países desenvolvidos, pois a população idosa está aumentando de forma alarmante, e os problemas de deglutição já estão entre as prioridades para identificação e monitoramento nas linhas de saúde.

[0027] A identificação da disfagia é uma tarefa extremamente importante devido às suas consequências nocivas para a saúde humana.

SUMÁRIO DA INVENÇÃO

[0028] A presente invenção apresenta um sistema de suporte para o diagnóstico de disfagia orofaríngea, o sistema compreendendo um ou mais processadores acoplados a pelo

menos um dispositivo de memória que armazena instruções executáveis por um ou mais processadores para pelo menos executar as operações de:

com um dispositivo sensor, adquirir sinais acústicos emitidos durante um evento de deglutição de um material por uma pluralidade de indivíduos;

identificar sinais acústicos de deglutição dentro dos sinais acústicos emitidos;

pré-processar os sinais acústicos de deglutição;

converter os sinais acústicos de deglutição em dados de imagem;

extrair características representativas do evento de deglutição a partir dos dados de imagem;

gerar um conjunto de dados compreendendo os dados de imagem;

dividir os dados de imagem em um conjunto de dados de treinamento e um conjunto de dados de teste;

inserir o conjunto de dados de treinamento em uma rede neural para treinar um modelo para classificar os dados de imagem com base em pelo menos uma das características, como indicativo de deglutição normal e deglutição anormal devido a disfagia orofaríngea;

validar o diagnóstico de disfagia orofaríngea usando dados de teste; e

emitir um modelo treinado para pelo menos diagnosticar

disfagia orofaríngea.

[**0029**] A presente invenção também apresenta um método para diagnosticar disfagia orofaríngea, o método compreendendo:

com um ou mais processadores acoplados a pelo menos um dispositivo de memória armazenando instruções de programa, um ou mais processadores executando as instruções de programa para executar as operações de:

receber sinais acústicos emitidos durante um evento de deglutição de um material por pelo menos um indivíduo, em que os sinais acústicos são adquiridos por um dispositivo sensor;

identificar sinais acústicos de deglutição dentro dos sinais acústicos emitidos;

pré-processar os sinais acústicos de deglutição;

converter os sinais acústicos de deglutição em dados de imagem;

extrair características representativas da deglutição dos dados de imagem;

gerar um conjunto de dados compreendendo os dados de imagem;

dividir os dados de imagem em um conjunto de dados de treinamento e um conjunto de dados de teste;

inserir o conjunto de dados de treinamento em uma rede neural para treinar um modelo para classificar os dados de imagem com base em pelo menos uma das características, como

indicativo de deglutição normal e deglutição anormal devido a disfagia orofaríngea;

validar o diagnóstico de disfagia orofaríngea usando dados de teste; e

emitir um modelo treinado para pelo menos diagnosticar disfagia orofaríngea.

[0030] Em outra modalidade, a presente invenção divulga um método para diagnóstico de disfagia orofaríngea compreendendo operações realizadas por um aparelho compreendendo um ou mais processadores acoplados a pelo menos um dispositivo de memória armazenando instruções executáveis por um ou mais processadores para pelo menos executar as operações de:

receber um conjunto de dados de imagem associado à deglutição de sinais acústicos adquiridos de um paciente em questão engolindo um material;

aplicar o conjunto de dados de imagem a um modelo treinado para determinar a presença de pelo menos uma das características representativas do evento de deglutição dentro do conjunto de dados de imagem, como indicativo de pelo menos uma deglutição normal e deglutição anormal devido a disfagia orofaríngea; e

produzir um relatório compreendendo pelo menos um do nível de gravidade da disfagia orofaríngea e um risco de aspiração laringotraqueal.

[0031] A presente invenção também divulga um meio legível por computador que armazena instruções para diagnosticar disfagia orofaríngea, sendo as instruções executáveis por um ou mais processadores para, ao menos, realizar as seguintes operações:

receber sinais acústicos emitidos durante um evento de deglutição de um material por um paciente em questão;

identificar sinais acústicos de deglutição dentro dos sinais acústicos emitidos; pré-processar os sinais acústicos de deglutição;

converter os sinais acústicos de deglutição em dados de imagem;

extrair características representativas do evento de deglutição a partir dos dados de imagem; gerar um conjunto de dados compreendendo os dados de imagem;

dividir os dados de imagem em um conjunto de dados de treinamento e um conjunto de dados de teste;

inserir o conjunto de dados de treinamento em uma rede neural para treinar um modelo para classificar os dados de imagem com base em pelo menos uma das características, como indicativo de deglutição normal e deglutição anormal devido a disfagia orofaríngea;

validar o diagnóstico de disfagia orofaríngea usando dados de teste; e emitir um modelo treinado para pelo menos diagnosticar disfagia orofaríngea.

determinar um nível de gravidade da disfagia orofaríngea e prever um risco de aspiração laringotraqueal; e

produzir um relatório compreendendo pelo menos um do nível de gravidade da disfagia orofaríngea e um risco de aspiração laringotraqueal.

[0032] Em uma modalidade, a presente invenção divulga um modelo preditivo treinado para o diagnóstico de disfagia orofaríngea, o modelo preditivo treinado e configurado para pelo menos:

receber um conjunto de dados de imagem associado à deglutição de sinais acústicos adquiridos de um paciente em questão engolindo um material;

aplicar o conjunto de dados de imagem a um modelo treinado para determinar a presença de pelo menos uma das características representativas do evento de deglutição dentro do conjunto de dados de imagem, como indicativo de pelo menos uma deglutição normal e deglutição anormal devido a disfagia orofaríngea; e

produzir um relatório compreendendo pelo menos um do nível de gravidade da disfagia orofaríngea e um risco de aspiração laringotraqueal.

Solução do problema

[0033] É fornecido um sistema que compreende uma máquina com uma rede neural que usa algoritmos específicos para aumentar

a precisão diagnóstica da disfagia orofaríngea.

[0034] A presente divulgação descreve um sistema de suporte de diagnóstico compreendendo pré-processamento e análise de sons de deglutição e sua classificação através de redes neurais de aprendizagem profunda, indicando o grau de disfagia do indivíduo. O sistema de suporte diagnóstico permite a padronização da avaliação clínica da deglutição em uma população e fornece resultados diagnósticos reprodutíveis e precisos. Este sistema pode ser utilizado em qualquer lugar e por qualquer profissional e pode ser utilizado na avaliação e diagnóstico da deglutição.

[0035] A presente divulgação descreve um sistema de suporte diagnóstico prático e de baixo custo que pode avaliar adequadamente todo o processo de deglutição em diferentes indivíduos, permitindo assim um menor número de consultas, maior resolução nas avaliações, detecção precoce de identificação, padronização dos achados, pesquisa confiável. Além disso, o sistema de suporte diagnóstico pode ser usado para avaliar a mastigação por fonoaudiólogos e dentistas, para verificar a adaptação de próteses e procedimentos pós-cirúrgicos, para treinar a mastigação adequada e para reabilitação mastigatória com deglutição adequada.

[0036] Além disso, os métodos e o sistema aqui descritos compreendem redes neurais que podem ser continuamente treinadas para melhorar os resultados, minimizando os erros humanos e, assim, acelerar a avaliação da deglutição. Como tal, o uso de inteligência artificial melhora a triagem, o

diagnóstico e o monitoramento clínico preciso de indivíduos com disfagia, reduzindo os custos de saúde, expandindo o acesso à reabilitação adequada e reduzindo os problemas de saúde e a mortalidade resultantes da disfagia.

[0037] Como exemplo, um profissional realiza a ausculta cervical no paciente enquanto o paciente está engolindo e registra esses sons. O sistema de suporte de diagnóstico usa os sons gravados para fornecer a um paciente um diagnóstico instantâneo pelo profissional de saúde com base nos resultados do sistema de suporte de diagnóstico. Além disso, o paciente pode receber orientação especializada e sugestões específicas de encaminhamento, tudo a partir da análise acústica da deglutição. As análises resultantes do processamento são personalizadas e individualizadas para cada tipo de problema e paciente, considerando todos os dados clínicos fornecidos quando o paciente foi cadastrado no sistema e posterior análise dos sons da deglutição. Então, usando um processador central, um histórico personalizado é criado para monitoramento subsequente do paciente, que é meticulosamente estudado pelo programa com base em parâmetros acústicos contidos em seu banco de dados. Usando as informações coletadas pelo sistema, é possível identificar possíveis disfagias, bem como seu grau de gravidade, receber suporte para o manejo clínico e monitorar seu prognóstico.

[0038] Em outro exemplo, os métodos e o sistema aqui descritos podem ser empregados por médicos, enfermeiros e

outros profissionais envolvidos na identificação precoce e reabilitação do processo de deglutição, incluindo em avaliações clínicas e avaliações através de videofluoroscopia da deglutição ou nasofibroendoscopia. Vantajosamente, os métodos e sistema aqui descritos melhoram a qualidade do diagnóstico clínico na área que está sendo oferecida por diferentes profissionais, independente da experiência e treinamento, além de permitir uma ação baseada em evidências.

BREVE DESCRIÇÃO DOS DESENHOS

[0039] A Figura 1 ilustra um sistema de exemplo que pode ser usado para realizar uma ou mais técnicas para o diagnóstico de disfagia orofaríngea;

[0040] A Figura 2 mostra um exemplo de rede neural de convolução (CNN - *convolution neural network*);

[0041] A Figura 3 mostra um exemplo de diagrama de fluxo representando um método implementado por computador para determinar as características associadas à disfagia;

[0042] A Figura 4a mostra um exemplo de sons de deglutição obtidos durante um evento de deglutição;

[0043] A Figura 4b mostra eventos sonoros relevantes dentro dos sons de deglutição;

[0044] A Figura 4c mostra um sinal acústico associado à deglutição em um domínio do tempo;

[0045] A Figura 4d mostra um exemplo de um espectrograma da

janela que engole o sinal acústico da Figura 4c;

[0046] A Figura 4e mostra um espectro de potência correspondente do sinal acústico da Figura 4c;

[0047] A Figura 4f mostra um periodograma correspondente do sinal acústico da Figura 4c;

[0048] A Figura 5a mostra um exemplo de uma primeira rede neural convolucional para realizar uma ou mais técnicas para diagnóstico de disfagia;

[0049] A Figura 5b mostra um exemplo de uma segunda rede neural convolucional para realizar uma ou mais técnicas para o diagnóstico de disfagia;

[0050] A Figura 6 mostra um exemplo de uma terceira rede neural convolucional compreendendo um modelo de conjunto para realizar uma ou mais técnicas para diagnóstico de disfagia;

[0051] As Figuras 7a e 7b mostram curvas de aprendizagem geradas a partir de experiências que compreendem 1.175 espectrogramas com sinais de deglutição de indivíduos saudáveis e 1.187 espectrogramas de indivíduos com disfagia; e

[0052] A Figura 8 mostra um diagrama de blocos de um exemplo de uma máquina sobre a qual qualquer uma ou mais das técnicas (por exemplo, metodologias) aqui discutidas podem ser realizadas.

DESCRIÇÃO DETALHADA DA INVENÇÃO

[0053] A seguinte descrição detalhada refere-se aos desenhos anexos. Sempre que possível, números de referência iguais ou semelhantes são usados nos desenhos e na descrição a seguir para se referir a partes ou componentes iguais ou semelhantes. Embora diferentes modalidades da divulgação possam ser descritas, modificações, adaptações e outras implementações são possíveis. Por exemplo, substituições, adições ou modificações podem ser feitas nas partes e componentes ilustrados nos desenhos, e os métodos ilustrativos descritos neste documento podem ser modificados substituindo, reordenando, removendo ou adicionando etapas aos métodos divulgados. Por conseguinte, a seguinte descrição detalhada não limita a divulgação. Em vez disso, o escopo adequado é definido pelas reivindicações anexas.

[0054] Além disso, deve ser apreciado que as implementações particulares mostradas e descritas neste documento são ilustrativas da invenção e não se destinam a limitar o escopo da invenção de qualquer forma. De fato, por uma questão de brevidade, certos subcomponentes dos componentes operacionais individuais e outros aspectos funcionais dos sistemas podem não ser descritos em detalhes neste documento. Além disso, as linhas de conexão mostradas nas várias figuras contidas neste documento têm a intenção de representar relações funcionais e/ou acoplamentos físicos exemplificativos entre os vários elementos. Deve ser notado que muitas relações funcionais ou conexões físicas alternativas ou adicionais podem estar presentes em um

sistema prático.

[0055] A Figura 1 ilustra a implementação de um sistema (10) de exemplo que pode ser usado para realizar uma ou mais técnicas para o diagnóstico de disfagia orofaríngea descrita nesta divulgação. Um dispositivo de diagnóstico (12), configurado para adquirir sinais acústicos ou sinais vibracionais emitidos quando um paciente em questão engole, é acoplado a uma máquina ou aparelho (14), tal como um dispositivo cliente. O dispositivo de diagnóstico (12), ou dispositivo sensor, é colocado em um paciente em questão (16) em uma área de garganta (18) e compreende um transdutor de áudio para ausculta, tal como gravação de sons de deglutição emitidos enquanto o paciente em questão engole.

[0056] Em um exemplo, o dispositivo de diagnóstico (12) é um estetoscópio digital, tal como um estetoscópio Littmann (Classic III) com um amplificador (Eko Core), da 3M, Minnesota, EUA.

[0057] A máquina (14), tal como um dispositivo de computação, compreende um ou mais processadores (20), um dispositivo de memória (22) para armazenar instruções legíveis por computador (24) executáveis pelo processador (20) e para armazenar dados associados aos sinais acústicos ou sinais vibracionais recebidos do dispositivo sensor (12).

[0058] Os sinais acústicos adquiridos ou sinais vibracionais são convertidos em dados acústicos ou dados vibracionais (26) pelo processador (20) usando instruções de programa

(24) e a máquina (14) usa instruções de programa (24), tais como algoritmos, para analisar os dados de imagem ou espectrogramas derivados dos sinais acústicos adquiridos para determinar a probabilidade de disfagia e auxiliar um clínico na prestação de um diagnóstico.

[0059] As instruções executáveis por máquina (24) compreendem redes neurais (25) com algoritmos para modelos de dados e modelos de processo e dados acústicos adquiridos (26), incluindo outros dados. Como tal, o processador (20) pode ser incorporado como um executor de instruções de software (24), em que as instruções de software (24) podem configurar especificamente o processador (20) para executar algoritmos e/ou operações aqui descritas quando as instruções de software (24) são executadas.

[0060] Alternativamente, o processador (20) pode executar funcionalidades codificadas. A máquina (14) também compreende um módulo de interface de entrada/saída (28) e um link de comunicação (30) interconectando os vários componentes (20, 22, 28) da máquina (14). A máquina (14) também compreende uma interface gráfica de usuário (GUI) (31) e um ou mais bancos de dados (32a, 32b) com registros de pacientes e dados de diagnóstico associados e dados de treinamento.

[0061] O dispositivo sensor (12) comunica através de uma ligação com fios ou ligação de dados sem fios (33) com a máquina (14). Em um exemplo, a máquina (14) é um dispositivo móvel de computação sem fio, como um smartphone (por

exemplo, Apple iPhone™), tablet (por exemplo, Apple iPad™) ou laptop.

[0062] A máquina (14) também pode ser acoplada comunicativamente a uma máquina de servidor (34), tal como um computador de servidor, através de uma rede de comunicação (36). Acoplado à máquina de servidor (34) está um ou mais bancos de dados (38), diretamente ou através da rede (36). Semelhante ao dispositivo cliente (14), a máquina de servidor (34) compreende um ou mais processadores (40), o dispositivo de memória (42) armazenando dados (44), incluindo modelos de dados e modelos de processo, e instruções (46), módulo de interface de I/O (48), interconectados pelo barramento de comunicações (50).

[0063] Em um exemplo, um mecanismo de rede neural profunda de exemplo (25) usa armazenamento de coeficiente configurável na forma de um banco de dados de coeficiente. Quando combinado dentro de um sistema de computador de exemplo (14), o mecanismo de rede neural é configurado pelo processador (20) selecionando coeficientes lidos de um banco de dados de coeficientes e preenchendo a rede neural com esses coeficientes. Coeficientes exemplificativos compreendem parâmetros incluindo pesos e vieses associados a neurônios individuais ou blocos de uma configuração de rede neural particular e, em algumas modalidades, incluem a especificação de uma ou mais funções de ativação e funções de transferência.

[0064] O mecanismo de rede neural (25) pode incluir

configurações de rede neural de diferentes tipos. Em um exemplo, o mecanismo de rede neural (25) inclui uma rede neural de feedforward, por exemplo, uma rede neural de propagação de feedforward. Em outras modalidades, o mecanismo de rede neural (25) inclui uma ou mais configurações de rede neural adicionais ou alternativas, sozinhas ou em combinação, incluindo, mas não se limitando a, redes neurais recorrentes (RNN), por exemplo, RNN de memória de curto prazo longa (LSTM), rede neural convolucional (CNN) ou outra configuração de gráfico acíclico direcionado de uma rede neural. Em outros exemplos, o mecanismo de rede neural (25) compreende uma configuração de rede neural com múltiplas camadas ocultas de diferentes tipos, por exemplo, uma ou mais camadas RNN e uma ou mais camadas CNN.

[0065] Como mostrado na Figura 2, em um exemplo, uma rede neural convolucional (CNN) (25), um tipo de rede neural profunda e feedforward, é utilizada para processar os espectrogramas (52), ou seja, dados de imagem, e extrair características aprendidas a partir dos dados acústicos adquiridos ou dos espectrogramas (52), com o objetivo de resolver o problema de classificação. O exemplo CNN (25) inclui várias camadas de convolução (54), camadas de agrupamento (56), uma camada de achatamento (57) e camadas totalmente conectadas (58) entre as camadas de entrada e saída.

[0066] A camada de convolução (54) inclui múltiplos filtros

dos pesos de rede. Além disso, a camada totalmente conectada (58) tem um grupo semântico de nós que são conectados através de bordas ponderadas aos nós nas camadas seguintes e anteriores. Encontrar o melhor conjunto de pesos para esses filtros de convolução e bordas de conexão desempenha um papel crucial no processo de treinamento da rede neural. Geralmente, a camada convolucional (54) converte a imagem de entrada (52) em um mapa de características, realizando uma operação linear chamada convolução.

[0067] O espectrograma (52) compreende uma matriz de valores de pixel e é uma matriz com três camadas diferentes. O tamanho da matriz depende da resolução e do tamanho da imagem, por exemplo, matriz de números 224 × 224 × 3, onde 3 se refere à cor RGB. Um valor de 0 a 255 é atribuído a cada componente da matriz para descrever a intensidade de cada ponto. Para extrair as diferenças entre várias imagens, um grupo de filtros é aplicado a essas imagens (52) e os filtros envolvem a entrada e passam o resultado para a próxima camada aplicando um produto escalar, convolução, sobre todo o tamanho de cada imagem (52). Os parâmetros desses filtros (isto é, os pesos) são inicializados aleatoriamente e precisam ser aprendidos pela rede neural (25) subsequentemente.

[0068] Após a camada convolucional (54), a camada de agrupamento (camada de pooling) (56) aplica as seguintes funções de agrupamento destinadas a reduzir o número de características nos mapas de características:

agrupamento médio, isto é, calcular a média de cada remendo da matriz de características; e

agrupamento máximo, isto é, calcular o valor máximo de cada remendo da matriz de características.

[0069] Após as etapas de convolução e agrupamento, toda a matriz de mapa de características obtida é transformada em um único vetor pela camada de achatamento (flattening) (57). Após o achatamento, o vetor de característica obtido é enviado para a camada totalmente conectada (fully connected layer) (58) como sua entrada. Este vetor de entrada é então conectado à camada de saída através de uma camada oculta totalmente conectada (58).

[0070] Por conseguinte, as camadas de convolução (54) e as camadas de agrupamento (56) atuam como extratores de recursos da imagem de entrada (52), enquanto as camadas totalmente conectadas (58) atuam como um classificador. Consequentemente, a taxa de erro de classificação será calculada usando as classes previstas e reais. O processo de treinamento recursivo continuará até encontrar a maior acurácia da classificação. Geralmente, o processo de treinamento envolve ajustar esses pesos dos filtros e neurônios e aplicá-los ao conjunto de dados disponível. Esses pesos são inicializados aleatoriamente no início e são atualizados usando diferentes técnicas de otimização em um processo recursivo.

[0071] Em um exemplo, foi utilizado um conjunto de dados

compreendendo 1.175 espectrogramas (52) com sinais de deglutição de indivíduos saudáveis e 1.187 espectrogramas (52) de indivíduos com disfagia. Um subconjunto do conjunto de dados foi usado para treinar o modelo durante a fase de treinamento e outro subconjunto foi usado para testar o modelo durante a fase de teste.

[0072] A estrutura de rede neural (25) é projetada para minimizar o sobreajuste, pelo que o classificador (58) não generaliza bem entre os dados de treinamento e os dados de teste, de modo que haja uma lacuna significativa entre a precisão do treinamento e a precisão do teste.

[0073] Em um exemplo, o modelo usa um otimizador Adam, um algoritmo de otimização iterativo usado para minimizar a função de perda durante o treinamento de redes neurais (25). Ao contrário de outros algoritmos de otimização que exigem o armazenamento de um histórico de gradientes para cada parâmetro, o otimizador Adam só precisa manter duas médias móveis por parâmetro, tornando-o assim eficiente em memória e computacionalmente eficiente.

[0074] A Figura 3 mostra um diagrama de fluxo (100) representando um método implementado por computador para determinar as características associadas a uma condição de disfagia processando dados acústicos ou dados vibracionais (26) de acordo com várias implementações aqui descritas. Deve ser entendido que, embora o diagrama de fluxo operacional indique uma ordem particular de execução das operações, em outras implementações, as operações podem ser

executadas em uma ordem diferente. Além disso, em algumas implementações, operações ou blocos adicionais podem ser adicionados ao método. Da mesma forma, algumas operações ou blocos podem ser omitidos.

[0075] Em um estudo, foi criado um banco de dados a partir de pesquisas epidemiológicas científicas aprovadas por um comitê de ética em pesquisa para o qual os participantes com suspeita de disfagia foram convidados a participar do estudo, bem como os respectivos controles sem suspeita. Indivíduos com 18 anos ou mais foram submetidos a um rigoroso protocolo de pesquisa no qual passaram pelas seguintes etapas: entrevista inicial, avaliação clínica com coleta acústica de sons de deglutição para saliva, líquidos, pastosos e sólidos controlados e livres, e qualquer combinação dos mesmos, incluindo videofluoroscopia de deglutição. A coleta de dados foi automatizada e seguiu um manual de campo. Os coletores eram profissionais treinados, calibrados e especialistas na área. O banco de dados foi composto por mais de 350 participantes e 5.600 eventos de deglutição.

[0076] Na etapa (102) do diagrama de fluxo (100), um dispositivo sensor (12), tal como um estetoscópio digital, ou outro dispositivo de gravação de som auscultatório é posicionado num paciente em questão.

[0077] Em um exemplo, uma avaliação da deglutição é iniciada solicitando ao paciente em questão que engula um material por um período predeterminado. O material pode incluir uma

saliva, um líquido, um bode expiatório e um sólido, ou qualquer combinação dos mesmos.

[0078] Na etapa (104) do diagrama de fluxo (100), enquanto o paciente em questão está engolindo, o estetoscópio digital (12) é mantido firmemente no local da ausculta, como uma área da garganta (18) do paciente em questão (16), e a ausculta começa. Após o término do período predeterminado para deglutição e registro, o estetoscópio digital (12) é removido do local (18). Aos sons de deglutição adquiridos é atribuído um identificador único associado ao paciente em questão. Além disso, um identificador único pode estar associado ao nome do paciente, idade, data, comorbidades, dieta atual, avaliação ou reavaliação. Detalhes do clínico que realiza a avaliação também podem ser incluídos. Uma notação ou rótulo sobre o tipo de matéria que está sendo engolida, por exemplo, saliva, um líquido, um bode expiatório e um sólido. Os eventos de som de deglutição associados são convertidos em dados de evento de som auscultatório, como um arquivo de áudio digital (por exemplo, formato de arquivo .wav ou formato de arquivo .mp3) e armazenados no dispositivo de memória (22).

[0079] Na etapa (106) do diagrama de fluxo (100), os sinais acústicos (60) são pré-processados por condicionamento e subsequentemente passando através de um filtro de atenuação de ruído, que pode compreender um algoritmo de redução de ruído para remover ruído.

[0080] Na etapa (108) do diagrama de fluxo (100), como

mostrado na Figura 4b, a etapa de pré-processamento identifica os eventos sonoros relevantes em cada sinal acústico e extrai os trechos correspondentes aos eventos identificados (62, 64) em janelas de tamanho fixo predefinido (duração). Por conseguinte, um membro de janela segmenta cada sinal acústico em pedaços de dados tipicamente em mudanças no padrão de sinal, como ilustrado por linhas verticais pontilhadas na Figura 4b.

[0081] A Figura 4c mostra um sinal acústico (70) associado à deglutição em um domínio do tempo. O sinal acústico (60) é normalizado e transformado em espectrogramas (52) usando transformada de Fourier de Tempo Curto (STFT), por exemplo, a Figura 4d mostra um exemplo de um espectrograma (52) da janela engolindo o sinal de áudio da Figura 4c. Geralmente, o espectrograma (52) mostra como o conteúdo de frequência do sinal muda ao longo do tempo e pode ser calculado a partir do sinal no domínio do tempo.

[0082] Mais detalhadamente, os segmentos acústicos ou fragmentos de dados resultantes são inseridos em um módulo de extração de recursos. Para cada segmento, o módulo de extração de característica extrai o padrão de sinal de interesse do segmento/fragmento de dados e produz pontos de dados de segmento acústico representativos da característica extraída ou vetores de característica.

[0083] Em um exemplo, critérios baseados em amplitude são usados para posicionar cada janela sobre cada evento e as janelas selecionadas associadas ao evento de deglutição são

identificadas e rotuladas como tal.

[0084] Os espectrogramas marcados (52) são armazenados num meio legível por computador como imagens, isto é, dados de imagem para formar um conjunto de dados para entrada na rede neural (25). A Figura 4e mostra um espectro de potência correspondente (72) do sinal acústico associado a um evento de deglutição e a Figura 4f mostra um periodograma correspondente (74) do sinal acústico associado a um evento de deglutição.

[0085] Na etapa (110) do diagrama de fluxo (100), o conjunto de dados acústicos marcado é dividido em dois subconjuntos, isto é, um conjunto de dados de treinamento para treinar uma rede neural (25) (etapa 112) e um conjunto de dados de teste (etapa 114) com objetos de dados para testar o classificador (58). Geralmente, subconjuntos de conjuntos de dados são selecionados aleatoriamente para treinar os modelos preditivos ou para avaliar os modelos preditivos, por exemplo, uma divisão 80:20.

[0086] Na etapa (116) do diagrama de fluxo (100), os dados de imagem de espectrograma marcados (52) são alimentados em uma primeira rede neural de aprendizagem profunda (25), que é configurada para classificar os sinais em uma de uma variedade de classes, tais como: sons de deglutição, tosse, eletrocardiograma (ECG), respiração e outros ruídos.

[0087] Em um exemplo, como mostrado na Figura 5a, a primeira rede neural convolucional (25) compreende:

pelo menos 4 modelos binários com uma imagem (espectrograma);

tamanho de entrada: 224.224,3;

1 camada convolucional (64);

3 camadas convolucionais (128.256.512), 2 Denso (fully connected layer) (512) e 1 Denso (fully connected layer) (1).

[0088] Um corpus de treinamento marcado ou dados de imagem de espectrograma (52) compreendendo uma coleção de sinais acústicos anotados conhecidos e previamente analisados anotados com classes de disfagia correspondentes também são alimentados na primeira rede neural de aprendizagem profunda (deep learning) (25). O processo de treinamento analisa os vetores de feições extraídos pelo processo de feições para selecionar e treinar o classificador (58).

[0089] Em um exemplo, diferentes classificadores ou algoritmos de aprendizagem de máquina são utilizados para construir os modelos preditivos, tais como algoritmos de aprendizagem supervisionada, algoritmos de aprendizagem não supervisionada e algoritmos de aprendizagem de reforço. Exemplos de sistemas de algoritmos de aprendizagem supervisionada incluem máquina de vetores de suporte, árvore de decisão, regressão linear, regressão logística, Bayes ingênuo, k-vizinhos mais próximo (KNN), floresta aleatória, AdaBoost, XGBoost e métodos de rede neural. Exemplos de sistemas de algoritmos de aprendizagem não supervisionados

incluem K-means, deslocamento médio, propagação de afinidade, agrupamento hierárquico, DBSCAN (agrupamento espacial baseado em densidade de aplicativos com ruído), modelo de mistura Gaussiana (GMM), campos aleatórios de Markov (CAM), isodata (dados de auto-organização iterativa) e sistemas C-means difusos. Exemplos de sistemas de algoritmos de aprendizagem por reforço incluem sistemas Maja e Teaching-Box. Geralmente, treinar os modelos preditivos envolve otimizar os parâmetros de um sistema preditivo para minimizar a função de perda.

[0090] Na etapa (118) do diagrama de fluxo (100), além da etapa de treinamento, os modelos preditivos também passam por validação usando conjuntos de dados de teste. Os modelos de rede de aprendizagem profunda construídos são inspirados em arquiteturas de rede pré-treinadas com base em métodos de dimensionamento composto altamente eficazes. Estes modelos são treinados com imagens obtidas a partir de espectrogramas, e o processo de treinamento é avaliado com base no método de validação cruzada estratificada.

[0091] Durante o processamento, cada classificador (58) em um conjunto de classificadores pode ser selecionado, desmarcado, modificado, testado, treinado ou modificado de outra forma. Isso inclui modificar o algoritmo que define o classificador (58), alterar os parâmetros ou condições do classificador (58) usados para treinar o classificador (58) e retreinar os classificadores candidatos (58) devido à disponibilidade de vetores de recursos adicionais ou à

modificação dos vetores de recursos disponíveis.

[0092] Na etapa (120) do diagrama de fluxo (100), a avaliação do modelo mede a capacidade de generalização do algoritmo através de métricas de desempenho, que avaliam o quão corretamente o modelo está classificando os sons do processo de deglutição. Métricas de avaliação como precisão, exatidão, sensibilidade e especificidade são usadas para analisar o desempenho da rede neural implementada (25).

[0093] Uma medida de acurácia para os classificadores candidatos (58) treinados pelo processo de treinamento é determinada e o modelo é avaliado através de feedback fornecido ao processo de recurso, ao processo de treinamento ou a ambos. Além disso, o feedback continua preferencialmente como um processo iterativo até que um critério de parada predeterminado seja atendido, isto é, até que o melhor modelo de CNN seja encontrado (etapa 122). Um subconjunto de características ou atributos para fazer as melhores previsões é selecionado.

[0094] Em um exemplo, cada recurso de entrada é multiplicado por algum valor, ou peso, de modo que durante a etapa de treinamento, os pesos sejam atualizados até que o melhor modelo seja encontrado.

[0095] Em um exemplo, o processo de treinamento compreende:

aprendizado supervisionado com validação cruzada de 5 vezes;

um Adam Optimizer; e

Perda: 0,001.

[0096] Os modelos binários e multiclasse são treinados separadamente, após o que são concatenados ao modelo final (ensemble). Os pesos estão associados a variáveis em cada camada de rede neural. Um peso pode ser atribuído a cada modelo de aprendizado de máquina incluído no modelo de conjunto.

[0097] Por exemplo, cada camada tem dois conjuntos de pesos:

peso número 1: este conjunto de pesos está associado às conexões entre as unidades da camada anterior (camada de entrada ou anterior) e as unidades da camada atual; e

peso número 2: este conjunto de pesos está associado às unidades da própria camada atual.

[0098] Olhando para a Tabela A e a Tabela B, abaixo, cada linha na saída corresponde a um conjunto de pesos numa camada de rede específica.

[0099] As tabelas abaixo informam como os pesos estão distribuídos na rede e quantos valores de peso estão em cada conjunto. Esses pesos são ajustados durante o treinamento em rede para aprender a representação de características e padrões relevantes nos dados de entrada. As formas dos pesos são definidas pelo projeto de rede e refletem a arquitetura específica para modelos binários (Tabela A) e multiclasse (Tabela B).

[0100] Tabela A

Nome da camada	Número do peso	Forma do peso
Conv2d	1	(3, 3, 3, 64)
Conv2d	2	(64,)
Conv2d_1	1	(3, 3, 64, 128)
Conv2d_1	2	(128,)
Conv2d_2	1	(3, 3, 128, 256)
Conv2d_2	2	(256,)
Conv2d_3	1	(3, 3, 256, 512)
Conv2d_3	2	(512,)
Dense	1	(100352, 512)
Dense	2	(512,)
Dense_1	1	(512, 512)
Dense_1	2	(512,)
Dense_2	1	(512, 1)
Dense_2	2	(1,)

[0101] Tabela B

Nome da camada	Número do peso	Forma do peso
Conv2d	1	(3, 3, 3, 64)
Conv2d	2	(64,)
Conv2d_1	1	(3, 3, 64, 128)
Conv2d_1	2	(128,)
Conv2d_2	1	(3, 3, 128, 256)
Conv2d_2	2	(256,)
Conv2d_3	1	(3, 3, 256, 512)
Conv2d_3	2	(512,)
Dense	1	(100352, 512)
Dense	2	(512,)
Dense_1	1	(512, 512)
Dense_1	2	(512,)
Dense_2	1	(512, 4)
Dense_2	2	(4,)

[0102] As Figuras 7a e 7b mostram as curvas de aprendizagem realizadas com uma consistência pastosa que foram geradas a partir de experimentos compreendendo 1.175 espectrogramas com sinais de deglutição de indivíduos saudáveis e 1.187 de indivíduos com disfagia. Como mostrado na Figura 5b, uma

segunda rede neural convolucional compreende:

um modelo multiclasse com um tamanho de entrada de imagem (espectrograma): 224.224,3;

1 camada convolucional (64);

3 camadas convolucionais (128.256.512), 2 densas (512) e 1 densa (4).

[0103] Para cada iteração do sistema (10), podem ser feitas alterações no conjunto de características candidato, nos classificadores candidatos (58) ou nos vetores de características extraídos com base no conjunto de características candidato e é determinada uma nova medida de desempenho.

[0104] Na etapa (124), através desta abordagem de feedback iterativo, um classificador robusto ou um modelo final de CNN pode ser gerado com base em um conjunto de treinamento mínimo e, de preferência, com mínima ou nenhuma intervenção humana.

[0105] Na etapa (126), os espectrogramas classificados como compreendendo sons de deglutição são avaliados por uma segunda rede neural de aprendizagem profunda (25) treinada para classificar os sons como sons de deglutição de indivíduos saudáveis ou indivíduos com disfagia.

[0106] Na etapa (128), os dados do teste de divisão são aplicados a uma terceira rede neural de aprendizagem profunda (25), que também recebe os sons classificados como

sons de deglutição de indivíduos disfágicos pela segunda rede neural de aprendizagem profunda (25).

[0107] A terceira rede neural de aprendizagem profunda (25) classifica então a saída da segunda rede neural de aprendizagem profunda (25) em diferentes níveis de gravidade da disfagia. Consequentemente, cada camada de nós na pluralidade de camadas de uma terceira rede neural de aprendizagem profunda (25) pode ser treinada em um conjunto distinto de características com base na saída da camada anterior para classificar sons de deglutição de diferentes níveis de disfagia.

[0108] Como mostrado na Figura 6, uma terceira rede neural convolucional compreende um modelo de conjunto, com uma entrada compreendendo os pesos de 4 modelos binários treinados e o modelo multiclasse reduz o número de predições de falsos positivos e aumenta o número de predições de verdadeiros positivos.

[0109] Em um exemplo, a função de perda usada para modelos de classificação binária é binary_crossentropy, enquanto para modelos de classificação multiclasse, é categorical_crossentropy. Após o treinamento e validação dos modelos preditivos nas etapas anteriores, os modelos preditivos treinados são emitidos e armazenados em um banco de dados de modelo treinado (etapa 130). Os modelos treinados são usados para prever disfagia quando fornecidos com dados de pacientes sujeitos não marcados (etapa 132). Como tal, os modelos preditivos treinados podem prever a

disfagia e a gravidade da disfagia.

[0110] Na etapa (134), é gerado um relatório de análise do processo de deglutição do paciente, contendo parâmetros biomédicos do sinal, introduzido pelo utilizador neste sistema para apoiar o diagnóstico de disfagia orofaríngea. Este relato compreende o sinal do processo de deglutição do paciente no domínio do tempo, destacando os sons da deglutição e a gravidade da disfagia do indivíduo. Além disso, o número de respirações, tosses, ECGs e outros ruídos, se houver, estão incluídos no relatório. Os dados acústicos da deglutição também podem ser descritos como amplitude, duração, magnitude do sinal e frequência (espectro de potência médio e frequência de pico) do som da deglutição. Com base nos resultados, o manejo clínico da disfagia pode ser sugerido a um paciente pelo profissional de saúde.

[0111] Em uma modalidade, os métodos descritos neste documento compreendem uma operação adicional de pelo menos realizar análise acústica dos sinais acústicos de deglutição para rastrear disfagia orofaríngea.

[0112] Em outra modalidade, os métodos descritos neste documento compreendem operação adicional de identificação de pelo menos uma alteração nos sinais acústicos de deglutição.

[0113] Em outra modalidade, os métodos descritos neste documento compreendem uma operação adicional de identificação de pelo menos uma alteração em um processo de

deglutição.

[0114] Em outra modalidade, os métodos descritos neste documento compreendem uma operação adicional de pelo menos prever um risco de aspiração laringotraqueal.

[0115] Em outra modalidade, os métodos descritos neste documento compreendem uma operação adicional de pelo menos realizar análise acústica dos sinais acústicos de deglutição para apoiar o diagnóstico em uso em uma aplicação de telemedicina.

[0116] Em outra modalidade, os métodos e sistema descritos neste documento podem ser implementados usando aprendizagem em conjunto, ou seja, combinando as previsões de múltiplos algoritmos de aprendizagem de máquina para melhorar o desempenho do modelo;

transferência de aprendizagem, ou seja, alavancando modelos pré-treinados e ajustando-os para a tarefa específica, muitas vezes melhorando o desempenho;

arquiteturas híbridas, ou seja, integrando diferentes tipos de redes neurais dentro de um único modelo, como combinar CNNs com LSTMs (Long Short-Term Memory); e camadas podem ser adicionadas ou reduzidas, e pesos, núcleos ou hiperparâmetros podem ser ajustados para os modelos.

[0117] Em outra modalidade, a precisão de diagnóstico do sistema de suporte (10) pode ser aumentada aumentando o tamanho da amostra. Além disso, o sistema pode ser usado com pessoas com diferentes patologias e características, tais

como o uso de uma traqueostomia. Quanto maior a heterogeneidade e possível extrapolação dos resultados, mais poderemos apoiar a tomada de decisão do profissional de saúde. Além disso, o sistema (10) pode ser configurado para atender a um determinado grupo demográfico, tal como crianças menores de 18 anos de idade.

[0118] Em uma implementação, a rede neural (25) pode ser implementada como sistema de hardware de computador independente ou como um módulo dentro de um sistema de computador maior sem se desviar do escopo da presente tecnologia.

[0119] Mais detalhadamente, a Figura 8 ilustra um diagrama de blocos de uma máquina de exemplo (200), tal como o dispositivo cliente (14) ou o computador servidor (34), sobre o qual qualquer uma ou mais das técnicas (por exemplo, metodologias) aqui discutidas podem ser realizadas.

[0120] A máquina (200) (por exemplo, um sistema de computador) pode incluir um processador de hardware (201) (por exemplo, uma unidade de processamento central (CPU), uma unidade de processamento gráfico (GPU), um núcleo de processador de hardware ou qualquer combinação dos mesmos), uma memória principal (202) e uma memória estática (206), conectada através de um interlink (203) (por exemplo, link ou barramento), uma vez que alguns ou todos esses componentes podem constituir hardware para sistemas ou implementações relacionadas discutidas acima.

[0121] De modo geral, o processador de hardware (201) pode, por exemplo, incluir pelo menos uma de uma Unidade de Processamento Central (CPU), um Processador de Computação de Conjunto de Instruções Reduzidas (RISC), um Processador de Computação de Conjunto de Instruções Complexas (CISC), uma Unidade de Processamento Gráfico (GPU), um Processador de Sinal Digital (DSP), uma Unidade de Processamento de Tensor (TPU), uma Unidade de Processamento Neural (NPU), uma Unidade de Processamento de Visão (VPU), um Acelerador de Aprendizado de Máquina, um Acelerador de Inteligência Artificial, um Circuito Integrado Específico de Aplicação (ASIC), uma Matriz de Portas Programáveis em Campo (FPGA), um Circuito Integrado de Radiofrequência (RFIC), um Processador Neuromórfico, um Processador Quântico ou qualquer combinação dos mesmos.

[0122] Um circuito de processador pode ainda ser um processador de múltiplos núcleos tendo dois ou mais processadores independentes (por vezes referidos como "núcleos") que podem executar instruções contemporaneamente. Os processadores de múltiplos núcleos contêm múltiplos núcleos computacionais em uma única matriz de circuito integrado, cada um dos quais pode executar o programa independentemente das instruções em paralelo.

[0123] O processamento paralelo em processadores de múltiplos núcleos pode ser implementado por meio de arquiteturas como superescalar, VLIW, processamento vetorial ou SIMD que permitem que cada núcleo execute fluxos de

instruções separados simultaneamente. Um circuito de processador pode ser emulado em software, rodando em um processador físico, como um processador virtual ou circuito virtual. O processador virtual pode se comportar como um processador independente, mas é implementado em software em vez de hardware.

[0124] Exemplos específicos de memória principal (202) incluem memória de acesso aleatório (RAM) e dispositivos de memória semicondutores, que podem incluir locais de armazenamento em semicondutores, tais como registradores. Exemplos específicos de memória estática (206) incluem memória não volátil, tal como dispositivos de memória semicondutores (por exemplo, memória somente de leitura eletricamente programável (EPROM), memória somente de leitura programável eletricamente apagável (EEPROM)) e dispositivos de memória flash; discos magnéticos, tais como discos rígidos internos e discos removíveis; discos magneto-ópticos; RAM; ou meios ópticos, tais como discos de CD-ROM e DVD-ROM.

[0125] A máquina (200) pode incluir um dispositivo de exibição (210), um dispositivo de entrada (212) (por exemplo, um teclado) e um dispositivo de navegação de interface de usuário (UI) (214) (por exemplo, um mouse). Em um exemplo, o dispositivo de exibição (210), o dispositivo de entrada (212) e o dispositivo de navegação UI (214) podem ser uma tela sensível ao toque.

[0126] A máquina (200) pode ainda incluir um dispositivo de

armazenamento em massa (216) (por exemplo, unidade de acionamento), um dispositivo de geração de sinal (218) (por exemplo, um alto-falante), um dispositivo de interface de rede (220).

[0127] A máquina 200 pode ainda incluir um controlador de saída (228), tal como uma conexão serial (por exemplo, barramento serial universal (USB), paralela ou outra conexão com fio ou sem fio (por exemplo, infravermelho (IR), comunicação de campo próximo (NFC), etc.) para comunicar ou controlar um ou mais dispositivos periféricos (por exemplo, uma impressora, leitor de cartão, etc.).

[0128] A máquina 200 pode ainda incluir mais um dispositivo de sensor (229) acoplado ao mesmo, tal como um estetoscópio digital.

[0129] O dispositivo de armazenamento em massa (216) pode compreender um meio legível por máquina (222) no qual é armazenado um ou mais conjuntos de estruturas de dados ou instruções (204) (por exemplo, software) que incorporam ou são utilizados por qualquer uma ou mais das técnicas ou funções aqui descritas. As instruções (204) também podem residir, completamente ou pelo menos parcialmente, dentro da memória principal (202), dentro da memória estática (206) ou dentro do processador de hardware (201) durante a execução da mesma pela máquina (200).

[0130] Em um exemplo, uma ou qualquer combinação do processador de hardware (201), da memória principal (202),

da memória estática (206) ou do dispositivo de armazenamento em massa (216) compreende um meio legível por máquina.

[0131] Exemplos específicos de meios legíveis por máquina incluem um ou mais de memória não volátil, tais como dispositivos de memória semicondutores (por exemplo, EPROM ou EEPROM) e dispositivos de memória flash; discos magnéticos, tais como discos rígidos internos e discos removíveis; discos magneto-ópticos; RAM; ou meios ópticos, tais como discos de CD-ROM e DVD-ROM. Embora o meio legível por máquina seja ilustrado como um único meio, o termo "meio legível por máquina" pode incluir um único meio ou múltiplos meios (por exemplo, um banco de dados centralizado ou distribuído, ou caches e servidores associados) configurados para armazenar uma ou mais instruções (204).

[0132] O termo "meio legível por máquina" inclui, por exemplo, qualquer meio que seja capaz de armazenar, codificar ou transportar instruções para execução pela máquina (200) e que faça com que a máquina (200) execute qualquer uma ou mais das técnicas da presente divulgação ou faça com que outro aparelho ou sistema execute qualquer uma ou mais das técnicas, ou que seja capaz de armazenar, codificar ou transportar estruturas de dados usadas por ou associadas a tais instruções. Exemplos de meios legíveis por máquina não limitativos incluem memórias de estado sólido, meios ópticos ou meios magnéticos.

[0133] Exemplos específicos de memória estática incluem memória não volátil, tal como dispositivos de memória

semicondutores (por exemplo, memória somente de leitura eletricamente programável (EPROM), memória somente de leitura programável eletricamente apagável (EEPROM)) e dispositivos de memória flash; discos magnéticos, tais como discos rígidos internos e discos removíveis; discos magneto-ópticos; RAM; ou meios ópticos, tais como discos de CD-ROM e DVD-ROM.

[0134] Em alguns exemplos, os meios legíveis por máquina incluem meios legíveis por máquina não transitórios. Em alguns exemplos, os meios legíveis por máquina incluem meios legíveis por máquina que não são um sinal de propagação transitório.

[0135] As instruções (204) podem ser transmitidas ou recebidas, por exemplo, através de uma rede de comunicações (205) utilizando um meio de transmissão através do dispositivo de interface de rede (220) utilizando qualquer um de um número de protocolos de transferência (por exemplo, frame relay, protocolo de internet (IP), protocolo de controle de transmissão (TCP), protocolo de datagrama do usuário (UDP), protocolo de transferência de hipertexto (HTTP), etc.). Exemplos de redes de comunicação incluem uma rede de área local (LAN), uma rede de área ampla (WAN), uma rede de dados de pacotes (por exemplo, a Internet), redes de telefonia móvel (por exemplo, redes celulares), redes de telefone antigo simples (POTS) e redes de dados sem fio (por exemplo, Instituto de Engenheiros Elétricos e Eletrônicos (IEEE) 802.11 família de padrões conhecida como Wi-Fi®),

família de padrões IEEE 802.15.4, uma família de padrões Long Term Evolution (LTE) 4G ou 5G, uma família de padrões Universal Mobile Telecommunications System (UMTS), redes peer-to-peer (P2P), redes de comunicação por satélite, entre outros.

[0136] Em um exemplo, o dispositivo de interface de rede (220) inclui um ou mais conectores físicos (por exemplo, Ethernet, coaxial ou outra interconexão) ou uma ou mais antenas para acessar a rede de comunicações (205). Em um exemplo, o dispositivo de interface de rede (220) inclui uma ou mais antenas para se comunicar sem fio usando pelo menos uma das técnicas de saída múltipla de entrada única (SIMO), saída múltipla de entrada múltipla (mimo) ou saída única de entrada múltipla (MISO). Em alguns exemplos, o dispositivo de interface de rede (220) se comunica sem fio usando técnicas mimo de Múltiplos Usuários.

[0137] O termo "meio de transmissão" deve ser considerado como incluindo qualquer meio intangível que seja capaz de armazenar, codificar ou transportar instruções para execução pela máquina (200) e inclui sinais de comunicação digitais ou analógicos ou outro meio intangível para facilitar a comunicação de tal software.

[0138] Exemplos, como aqui descritos, podem incluir, ou podem operar em lógica ou um número de componentes, módulos ou mecanismos (todos referidos a seguir como "módulos"). Os módulos são entidades tangíveis (por exemplo, hardware) capazes de realizar operações especificadas e são

configurados ou dispostos de uma certa maneira. Em um exemplo, os circuitos são dispostos (por exemplo, internamente ou em relação a entidades externas, como outros circuitos) de uma maneira especificada como um módulo. Em um exemplo, a totalidade ou parte de um ou mais sistemas de computador (por exemplo, um sistema de computador autônomo, cliente ou servidor) ou um ou mais processadores de hardware são configurados por firmware ou software (por exemplo, instruções, uma porção de aplicativo ou um aplicativo) como um módulo que opera para executar operações específicas. Em um exemplo, o software pode residir em um meio de armazenamento legível por computador não transitório ou outro meio legível por máquina. Em um exemplo, o software, quando executado pelo hardware subjacente do módulo, faz com que o hardware execute as operações especificadas.

[0139] Por conseguinte, o termo "módulo" é entendido como abrangendo uma entidade tangível, seja uma entidade que é fisicamente construída, especificamente configurada (por exemplo, com fio) ou temporariamente (por exemplo, transitoriamente) configurada (por exemplo, programada) para operar de uma maneira especificada ou para executar parte ou a totalidade de qualquer operação aqui descrita. Considerando exemplos em que os módulos são configurados temporariamente, cada um dos módulos não precisa ser instanciado em nenhum momento no tempo.

[0140] Por exemplo, quando módulos de hardware compreendem um processador de uso geral configurado usando software, o

processador de uso geral pode ser configurado como respectivos módulos de hardware diferentes em tempos diferentes. Software pode, consequentemente, configurar um processador, por exemplo, para constituir um módulo de hardware particular em uma instância de tempo e constituir um módulo de hardware diferente em uma diferente instância de tempo.

[0141] As modalidades da matéria e as operações funcionais descritas neste relatório descritivo podem ser implementadas em circuitos eletrônicos digitais, em software ou firmware de computador tangivelmente incorporados, em hardware de computador, incluindo as estruturas divulgadas nesta especificação e seus equivalentes estruturais, ou em combinações de um ou mais deles.

[0142] As modalidades da matéria descrita neste relatório descritivo podem ser implementadas como um ou mais programas de computador, isto é, um ou mais módulos de instruções de programa de computador codificados em um meio de armazenamento não transitório tangível para execução por, ou para controlar a operação de, aparelhos de processamento de dados.

[0143] Alternativamente ou adicionalmente, as instruções do programa podem ser codificadas em um sinal propagado gerado artificialmente, por exemplo, um sinal elétrico, óptico ou eletromagnético gerado por máquina que é gerado para codificar informações para transmissão para um aparelho receptor adequado para execução por um aparelho de

processamento de dados. O meio de armazenamento de computador pode ser um dispositivo de armazenamento legível por máquina, um substrato de armazenamento legível por máquina, um dispositivo de memória de acesso aleatório ou serial ou uma combinação de um ou mais deles.

[0144] Um programa de computador que também pode ser referido ou descrito como um programa, software, um aplicativo de software, um app, um módulo, um módulo de software, um script ou código) pode ser escrito em qualquer forma de linguagem de programação, incluindo linguagens compiladas ou interpretadas, ou linguagens declarativas ou processuais, e pode ser implantado em qualquer forma, incluindo como um programa autônomo ou como um módulo, componente, sub-rotina ou outra unidade adequada para uso em um ambiente de computação. Um programa de computador pode, mas não precisa, corresponder a um arquivo em um sistema de arquivos. Um programa pode ser armazenado em uma porção de um arquivo que contém outros programas ou dados, por exemplo, um ou mais scripts armazenados em um documento de linguagem de marcação, em um único arquivo dedicado para o programa em questão, ou em múltiplos arquivos coordenados, por exemplo, arquivos que armazenam um ou mais módulos, subprogramas ou porções do código. Um programa de computador pode ser implementado para ser executado em um computador ou em múltiplos computadores que estão localizados em um local ou distribuídos através de múltiplos locais e interligados por uma rede de comunicação.

[0145] Embora porções dos programas ilustrados nas várias figuras sejam mostradas como módulos individuais que implementam os vários recursos e funcionalidades através de vários objetos, métodos ou outros processos, os programas podem, em vez disso, incluir uma série de submódulos, serviços de terceiros, componentes, bibliotecas e tais, conforme apropriado. Por outro lado, as características e funcionalidades de vários componentes podem ser combinadas em componentes únicos, conforme apropriado.

[0146] Os processos e fluxos lógicos descritos neste relatório descritivo podem ser realizados por um ou mais computadores programáveis executando um ou mais programas de computador para executar funções operando em dados de entrada e gerando saída. Os processos e fluxos lógicos também podem ser realizados por, e o aparelho também pode ser implementado como, circuitos lógicos de finalidade especial, por exemplo, uma CPU, uma GPU, um FPGA ou um ASIC.

[0147] O termo "interface gráfica de usuário" ou "GUI" pode ser usado no singular ou no plural para descrever uma ou mais interfaces gráficas de usuário e cada uma das exibições de uma interface gráfica de usuário específica. Portanto, uma GUI pode representar qualquer interface gráfica do usuário, incluindo, mas não se limitando a, um navegador da web, uma tela sensível ao toque ou uma interface de linha de comando (CLI) que processa informações e apresenta eficientemente os resultados das informações ao usuário. Em geral, uma GUI pode incluir uma pluralidade de elementos de

interface de usuário (UI), alguns ou todos associados a um navegador da Web, como campos interativos, listas suspensas e botões operáveis pelo usuário. Esses e outros elementos da interface do usuário podem estar relacionados ou representar as funções do navegador da web.

[0148] As modalidades do assunto descrito neste relatório descritivo podem ser implementadas em um sistema de computação que inclui um componente de back-end, por exemplo, como um servidor de dados, ou que inclui um componente de middleware, por exemplo, um servidor de aplicativos, ou que inclui um componente de front-end, por exemplo, um computador cliente com uma interface gráfica de usuário, um navegador da web ou um aplicativo por meio do qual um usuário pode interagir com uma implementação do assunto descrito nesta especificação, ou qualquer combinação de um ou mais desses componentes de back-end, middleware ou front-end.

[0149] Os componentes do sistema podem ser interconectados por qualquer forma ou meio de comunicação de dados digitais de cabo de aço e/ou sem fio, por exemplo, uma rede de comunicações (205). Geralmente, a máquina (200) pode ser um sistema de computação em um ambiente de computação em nuvem.

Notas

[0150] Cada um dos aspectos não limitativos neste documento pode ser independente ou pode ser combinado em várias permutações ou combinações com um ou mais dos outros

aspectos ou outro assunto descrito neste documento.

[0151] A descrição detalhada acima inclui referências aos desenhos anexos, que fazem parte da descrição detalhada. Os desenhos mostram, a título de ilustração, modalidades específicas nas quais a invenção pode ser praticada. Estas modalidades também são referidas geralmente como "exemplos". Tais exemplos podem incluir elementos além daqueles mostrados ou descritos. No entanto, os presentes inventores também contemplam exemplos em que apenas os elementos mostrados ou descritos são fornecidos.

[0152] Além disso, os presentes inventores também contemplam exemplos usando qualquer combinação ou permutação desses elementos mostrados ou descritos (ou um ou mais aspectos dos mesmos), seja em relação a um exemplo particular (ou um ou mais aspectos dos mesmos), ou em relação a outros exemplos (ou um ou mais aspectos dos mesmos) mostrados ou descritos neste documento.

[0153] Neste documento, os termos "um" ou "uma" são usados, como é comum em documentos de patentes, para incluir um ou mais de um, independentemente de quaisquer outras instâncias ou usos de "pelo menos um" ou "um ou mais".

[0154] Neste documento, o termo "ou" é usado para se referir a um ou não exclusivo, de modo que "A ou B" inclua "A, mas não B", "B, mas não A" e "A e B", a menos que indicado de outra forma.

[0155] Nas reivindicações anexas, os termos "incluindo" e

"nos quais" são usados como equivalentes em inglês simples dos respectivos termos "compreendendo" e "em que". Além disso, nas reivindicações a seguir, os termos "incluindo" e "compreendendo" são abertos, isto é, um sistema, dispositivo, artigo, composição, formulação ou processo que inclui elementos além daqueles listados após tal termo em uma reivindicação ainda são considerados como abrangidos pelo escopo dessa reivindicação.

[0156] Além disso, nas reivindicações a seguir, os termos "primeiro", "segundo" e "terceiro", etc. são usados apenas como rótulos, e não se destinam a impor requisitos numéricos a seus objetos.

[0157] Os exemplos de método aqui descritos podem ser implementados por máquina ou por computador, pelo menos em parte. Alguns exemplos podem incluir um meio legível por computador ou meio legível por máquina codificado com instruções operáveis para configurar um dispositivo eletrônico para executar métodos como descrito nos exemplos acima. Uma implementação de tais métodos pode incluir código, tal como microcódigo, código de linguagem de montagem, um código de linguagem de nível superior ou semelhantes. Tal código pode incluir instruções legíveis por computador para realizar vários métodos. O código pode formar porções de produtos de programa de computador. Tais instruções podem ser lidas e executadas por um ou mais processadores para permitir o desempenho de operações compreendendo um método, por exemplo. As instruções estão em

qualquer forma adequada, tal como, mas não se limitando a código fonte, código compilado, código interpretado, código executável, código estático, código dinâmico e semelhantes.

[0158] Além disso, em um exemplo, o código pode ser armazenado de forma tangível em um ou mais meios legíveis por computador tangíveis voláteis, não transitórios ou não voláteis, como durante a execução ou em outros momentos. Exemplos destes meios legíveis por computador tangíveis podem incluir, mas não estão limitados a, discos rígidos, discos magnéticos removíveis, discos ópticos removíveis (por exemplo, discos compactos e discos de vídeo digital), cassetes magnéticos, cartões de memória ou sticks, memórias de acesso aleatório (RAMs), memórias somente de leitura (ROMs) e semelhantes.

[0159] A descrição acima pretende ser ilustrativa e não restritiva. Por exemplo, os exemplos descritos acima (ou um ou mais aspectos dos mesmos) podem ser usados em combinação uns com os outros. Outras modalidades podem ser usadas, tal como por um versado na técnica ao revisar a descrição acima.

[0160] O Resumo é provido para permitir que o leitor determine rapidamente a natureza da divulgação técnica. Ele é apresentado com o entendimento de que ele não será usado para interpretar ou limitar o escopo ou o significado das reivindicações. Além disso, na Descrição Detalhada acima, várias características podem ser agrupadas para agilizar a divulgação. Isso não deve ser interpretado como pretendendo que uma característica divulgada não reivindicada seja

essencial para qualquer reivindicação. Em vez disso, a matéria inventiva pode estar em menos do que todas as características de uma modalidade divulgada particular.

[0161] Assim, as seguintes reivindicações são aqui incorporadas na Descrição Detalhada como exemplos ou modalidades, com cada reivindicação estando por conta própria como uma modalidade separada, e é contemplado que tais modalidades podem ser combinadas umas com as outras em várias combinações ou permutações. O âmbito da invenção deve ser determinado com referência às reivindicações anexas, juntamente com o alcance completo dos equivalentes aos quais tais reivindicações têm direito.

REIVINDICAÇÕES

1. Um sistema de suporte para o diagnóstico de disfagia orofaríngea **caracterizado por** compreender um ou mais processadores acoplados a pelo menos um dispositivo de memória armazenando instruções executáveis por um ou mais processadores para pelo menos executar as operações de:

com um dispositivo sensor, adquirir sinais acústicos emitidos durante um evento de deglutição de um material por uma pluralidade de indivíduos;

identificar sinais acústicos de deglutição dentro dos sinais acústicos emitidos;

pré-processar os sinais acústicos de deglutição;

converter os sinais acústicos de deglutição em dados de imagem;

extrair características representativas do evento de deglutição dos dados de imagem;

gerar um conjunto de dados compreendendo os dados de imagem;

dividir os dados de imagem em um conjunto de dados de treinamento e um conjunto de dados de teste;

inserir o conjunto de dados de treinamento em uma rede neural para treinar um modelo para classificar os dados de imagem com base em pelo menos uma das características, como indicativo de deglutição normal e deglutição anormal devido à disfagia orofaríngea;

validar o diagnóstico de disfagia orofaríngea usando dados de teste; e

produzir um modelo treinado para pelo menos diagnosticar disfagia orofaríngea.

2. Sistema de suporte, de acordo com a reivindicação 1, **caracterizado por** compreender uma operação adicional de identificação de pelo menos uma alteração nos sinais acústicos de deglutição e/ou em um processo de deglutição.

3. Sistema de suporte, de acordo com qualquer uma das reivindicações 1 a 2, **caracterizado por** compreender uma operação adicional de determinar um nível de gravidade da disfagia orofaríngea.

4. Sistema de suporte, de acordo com qualquer uma das reivindicações 1 a 2, **caracterizado por** compreender uma operação adicional de pelo menos prever um risco de aspiração laringotraqueal.

5. Sistema de suporte, de acordo com qualquer uma das reivindicações 1 a 4, **caracterizado por** compreender operações adicionais de:

receber dados de imagem do paciente sujeito associados a um evento de deglutição;

aplicar os dados de imagem do paciente sujeito ao modelo treinado; e

emitir um relatório compreendendo pelo menos um dos diagnósticos de disfagia orofaríngea, o nível de gravidade da disfagia orofaríngea, o risco de aspiração laringotraqueal e sugestões para tratamento clínico.

6. Sistema de suporte, de acordo com a reivindicação 1, **caracterizado por** compreender uma operação adicional de realizar análise acústica dos sinais acústicos de deglutição

para rastrear disfagia orofaríngea.

7. Sistema de suporte, de acordo com a reivindicação 1 e 6, **caracterizado por** compreender uma operação adicional de realizar análise acústica dos sinais acústicos de deglutição para dar suporte a um ou mais dos seguintes diagnósticos:

diagnóstico de disfagia orofaríngea com base na avaliação da deglutição;

diagnóstico de disfagia orofaríngea com base na videofluoroscopia da deglutição ou nasofibroendoscopia da deglutição;

diagnóstico em uso em uma aplicação de telemedicina.

8. Sistema de suporte, de acordo com a reivindicação 1, 6 e 7, **caracterizado por** compreender uma operação adicional de realizar análise acústica dos sinais acústicos de deglutição para um ou mais dos seguintes diagnósticos:

monitorar a deglutição;

avaliar a mastigação, considerada a fase oral da deglutição;

dar suporte à tomada de decisão sobre possível conduta clínica para um profissional de saúde;

verificar adaptação de próteses, procedimentos pós-cirúrgicos e treinamento para reabilitação mastigatória adequada.

9. Um método para diagnosticar disfagia orofaríngea **caracterizado por** compreender um ou mais processadores acoplados a pelo menos um dispositivo de memória armazenando instruções de programa, o um ou mais processadores executando

as instruções de programa para executar as operações de:

receber sinais acústicos emitidos durante um evento de deglutição de um material por pelo menos um indivíduo, em que os sinais acústicos são adquiridos por um dispositivo sensor;

identificar sinais acústicos de deglutição dentro dos sinais acústicos emitidos;

pré-processar os sinais acústicos de deglutição;

converter os sinais acústicos de deglutição em dados de imagem;

extrair características representativas da deglutição dos dados de imagem;

gerar um conjunto de dados compreendendo os dados de imagem;

dividir os dados de imagem em um conjunto de dados de treinamento e um conjunto de dados de teste;

inserir o conjunto de dados de treinamento em pelo menos uma rede neural para treinar um modelo preditivo para classificar os dados de imagem com base em pelo menos uma das características, como indicativo de deglutição normal e deglutição anormal devido à disfagia orofaríngea;

validar o diagnóstico de disfagia orofaríngea usando dados de teste; e

emitir um modelo preditivo treinado para pelo menos diagnosticar disfagia orofaríngea.

10. Método, de acordo com a reivindicação 9, **caracterizado**

pelo fato de que o dispositivo sensor é um estetoscópio digital colocado na área da garganta de pelo menos um indivíduo para ausculta.

11. Método, de acordo com a reivindicação 9 e 10, **caracterizado pelo** fato de que os sinais acústicos são adquiridos enquanto o paciente engole pelo menos um dos seguintes: saliva, um líquido, um pastoso, um sólido e qualquer combinação destes.

12. Método, de acordo com a reivindicação 9, 10 e 11, **caracterizado pelo** fato de que os sinais acústicos são classificados como pelo menos um dos seguintes: sons de deglutição, tosse, eletrocardiograma (ECG), respiração e outros ruídos.

13. Método, de acordo com a reivindicação 12, **caracterizado pelo** fato de que os sinais acústicos de deglutição são normalizados e transformados em espectrogramas.

14. Método, de acordo com qualquer uma das reivindicações 9 a 13, **caracterizado pelo** fato de que a pelo menos uma rede neural de aprendizagem é treinada para determinar pelo menos um dos níveis de gravidade da disfagia orofaríngea.

15. Método, de acordo com qualquer uma das reivindicações 9 a 14, **caracterizado pelo** fato de que a pelo menos uma rede neural de aprendizagem é treinada para prever um risco de aspiração laringotraqueal.

16. Um método para diagnóstico de disfagia orofaríngea **caracterizado por** compreender operações realizadas por um aparelho compreendendo um ou mais processadores acoplados a pelo menos um dispositivo de memória armazenando instruções

executáveis por um ou mais processadores para pelo menos executar as operações de:

receber um conjunto de dados de imagem associado a sinais acústicos de deglutição adquiridos de um paciente sujeito engolindo um material;

aplicar o conjunto de dados de imagem a um modelo treinado para determinar a presença de pelo menos uma das características representativas do evento de deglutição dentro do conjunto de dados de imagem, como indicativo de pelo menos uma de deglutição normal e deglutição anormal devido a disfagia orofaríngea; e

emitir um relatório compreendendo pelo menos um dos níveis de gravidade da disfagia orofaríngea e um risco de aspiração laringotraqueal.

17. Um método para diagnóstico de disfagia orofaríngea **caracterizado por** compreender um meio legível por computador que armazena instruções para diagnosticar disfagia orofaríngea, as instruções executáveis por um ou mais processadores para pelo menos realizar as operações de:

receber sinais acústicos emitidos durante um evento de deglutição de um material por um paciente sujeito;

identificar sinais acústicos de deglutição dentro dos sinais acústicos emitidos pré-processar os sinais acústicos de deglutição;

converter os sinais acústicos de deglutição em dados de imagem;

extrair características representativas do evento de deglutição dos dados de imagem;

gerar um conjunto de dados compreendendo os dados de imagem;

dividir os dados de imagem em um conjunto de dados de treinamento e um conjunto de dados de teste;

inserir o conjunto de dados de treinamento em pelo menos uma rede neural para treinar um modelo para classificar os dados de imagem com base em pelo menos uma das características, como indicativo de deglutição normal e deglutição anormal devido à disfagia orofaríngea;

validar o diagnóstico de disfagia orofaríngea usando dados de teste; e

emitir um modelo treinado para pelo menos diagnosticar disfagia orofaríngea, determinar um nível de gravidade da disfagia orofaríngea e prever um risco de aspiração laringotraqueal; e

emitir um relatório compreendendo pelo menos um dos níveis de gravidade da disfagia orofaríngea e o risco de aspiração laringotraqueal.

18. Um método para diagnóstico de disfagia orofaríngea **caracterizado por** compreender um modelo preditivo treinado para o diagnóstico de disfagia orofaríngea, o modelo preditivo treinado configurado para pelo menos:

receber um conjunto de dados de imagem associado a sinais acústicos de deglutição adquiridos de um paciente sujeito engolindo um material;

determinar a presença de pelo menos uma das características representativas do evento de deglutição dentro do conjunto de dados de imagem, como indicativo

de pelo menos uma de deglutição normal e deglutição anormal devido à disfagia orofaríngea; e

emitir um relatório compreendendo pelo menos um dos níveis de gravidade da disfagia orofaríngea e um risco de aspiração laringotraqueal.

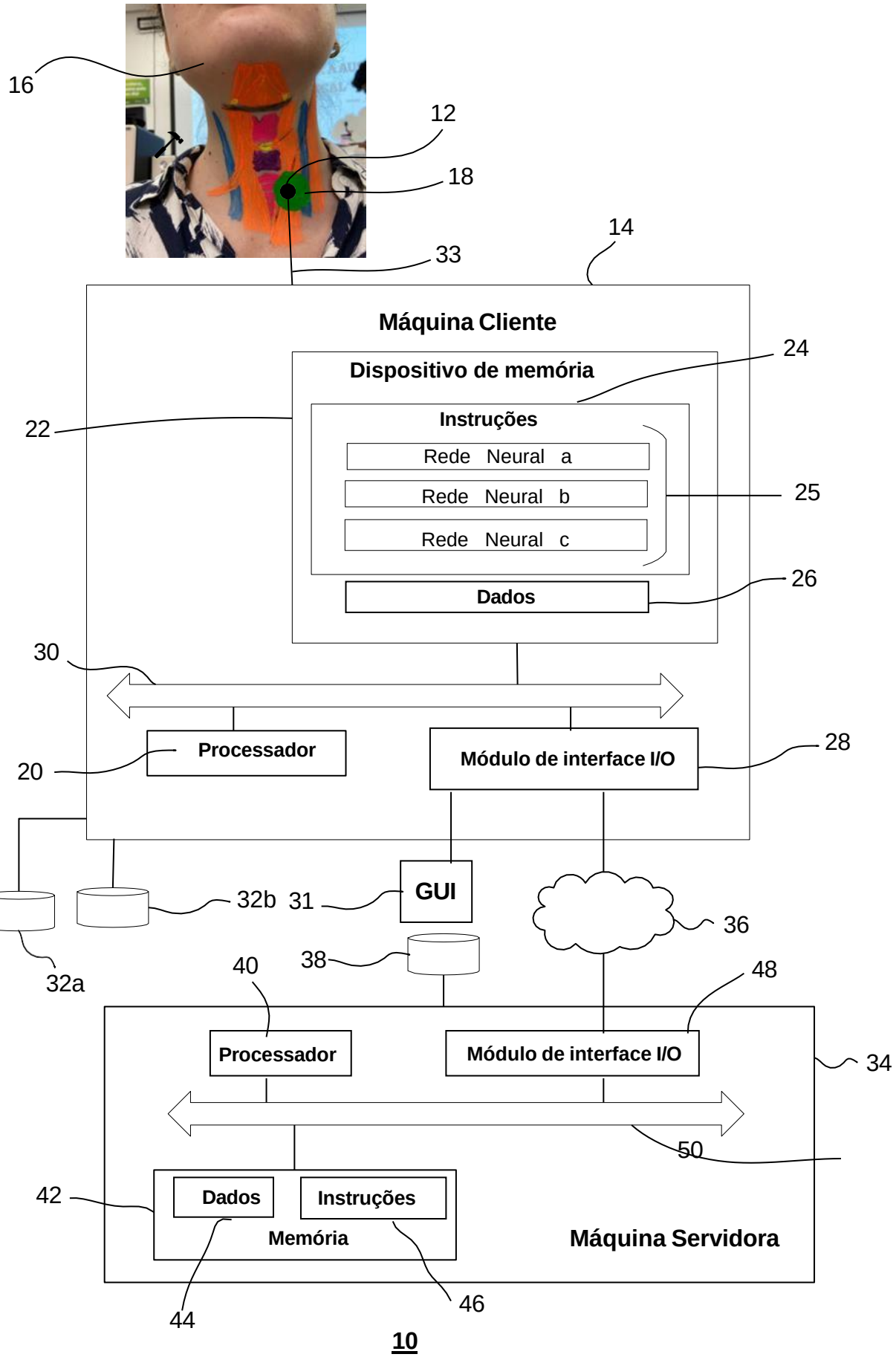

16
12
18
14
33
Máquina Cliente
24
Dispositivo de memória
22
Instruções
Rede Neural a
Rede Neural b
Rede Neural c
25
Dados
26
30
Processador
20
Módulo de interface I/O
28
GUI
31
32b
36
32a
38
40
48
Processador
Módulo de interface I/O
34
50
42
Dados
Instruções
Memória
Máquina Servidora
44
46
10

Figura 1

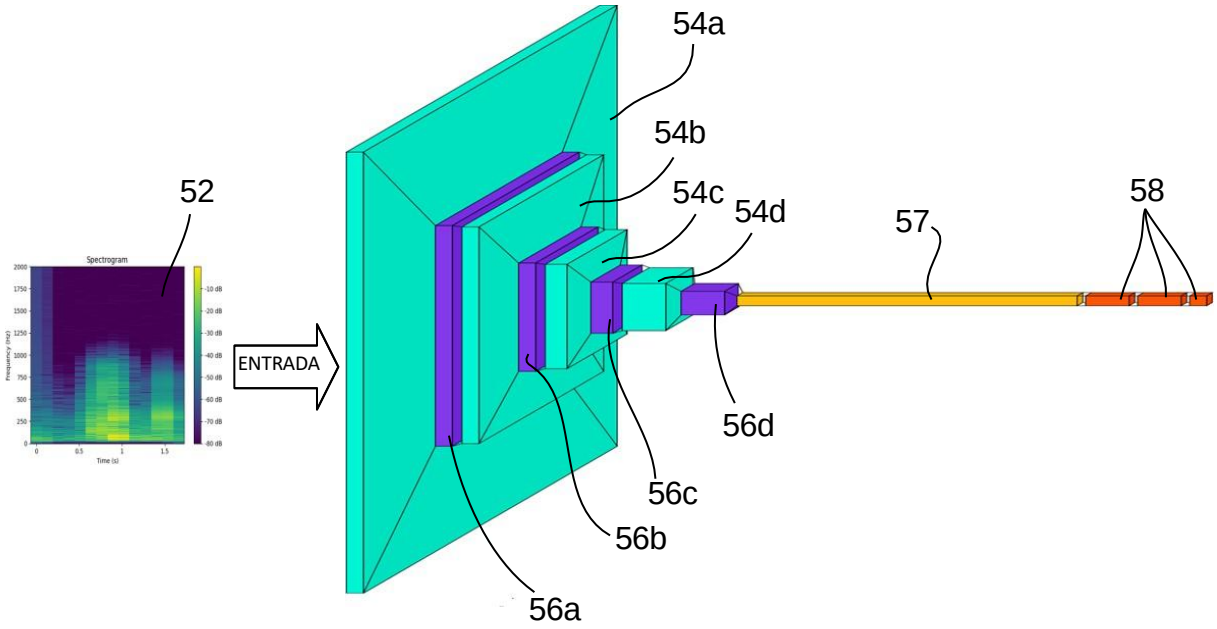

52
Spectrogram
ENTRADA
54a
54b
54c
54d
57
58
56d
56c
56b
56a
25

Figura 2

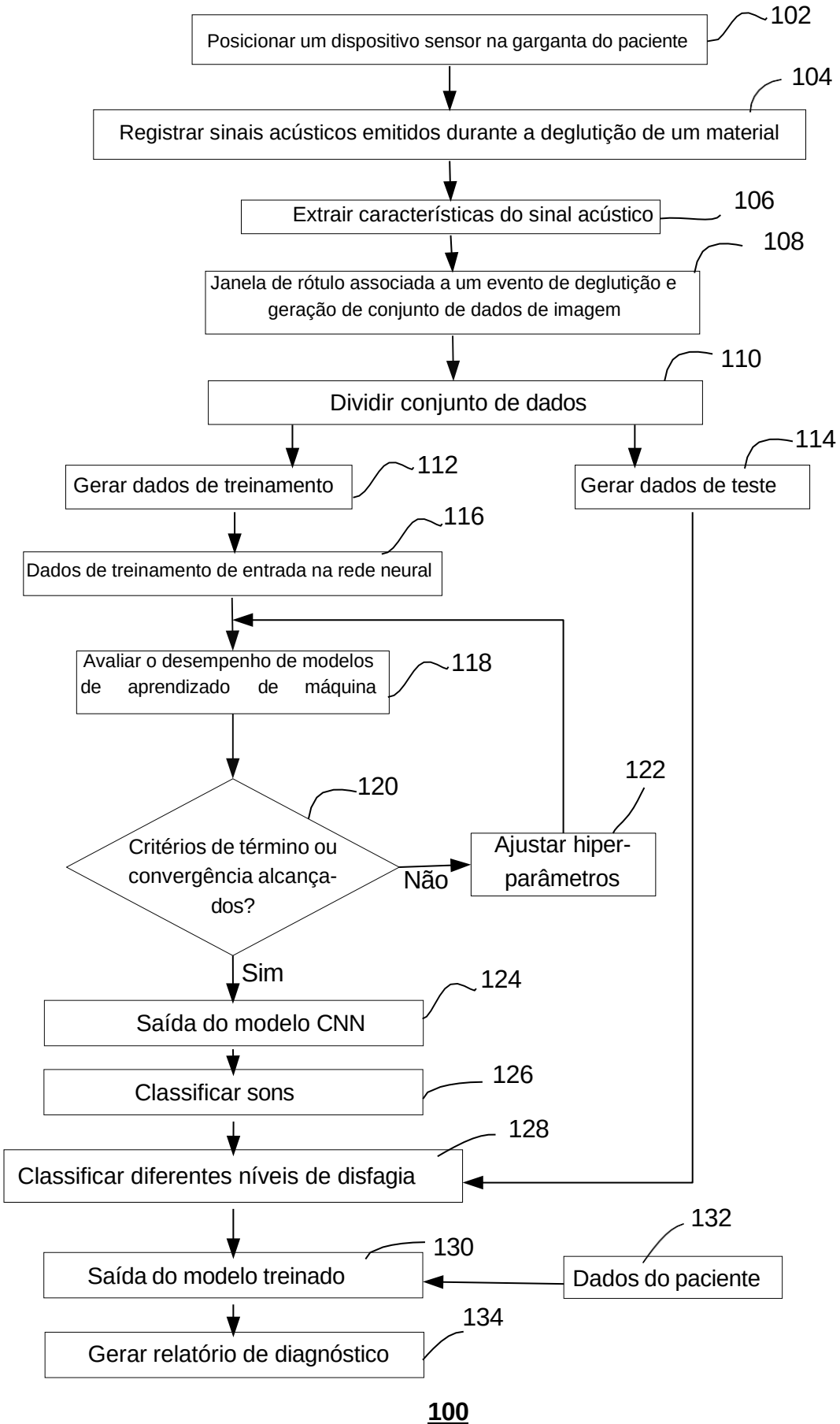

Posicionar um dispositivo sensor na garganta do paciente
102
Registrar sinais acústicos emitidos durante a deglutição de um material
104
Extrair características do sinal acústico
106
Janela de rótulo associada a um evento de deglutição e geração de conjunto de dados de imagem
108
Dividir conjunto de dados
110
Gerar dados de treinamento
112
Gerar dados de teste
114
Dados de treinamento de entrada na rede neural
116
Avaliar o desempenho de modelos de aprendizado de máquina
118
Critérios de término ou convergência alcançados?
120
Não
Ajustar hiperparâmetros
122
Sim
Saída do modelo CNN
124
Classificar sons
126
Classificar diferentes níveis de disfagia
128
Saída do modelo treinado
130
Dados do paciente
132
Gerar relatório de diagnóstico
134
100

Figura 3

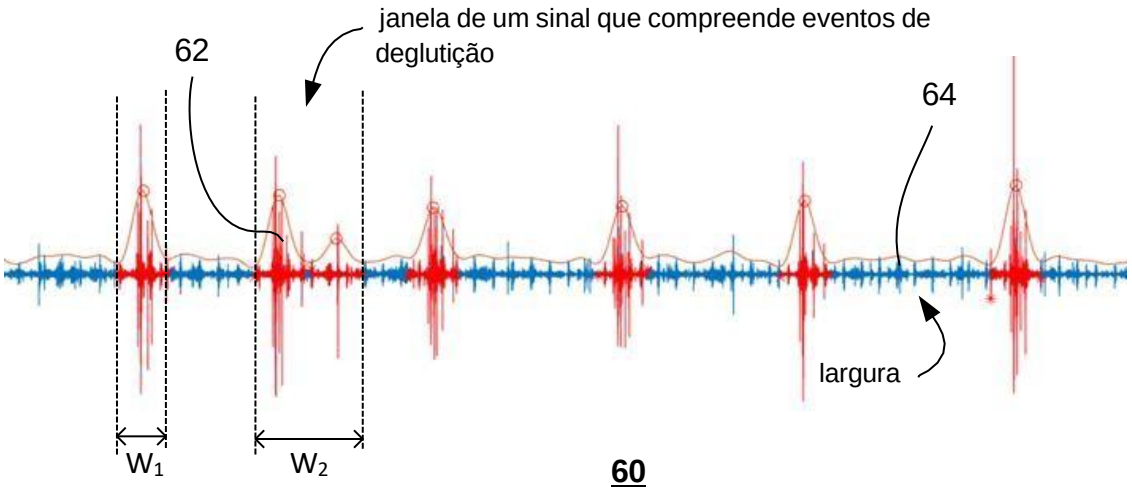


Figura 4a

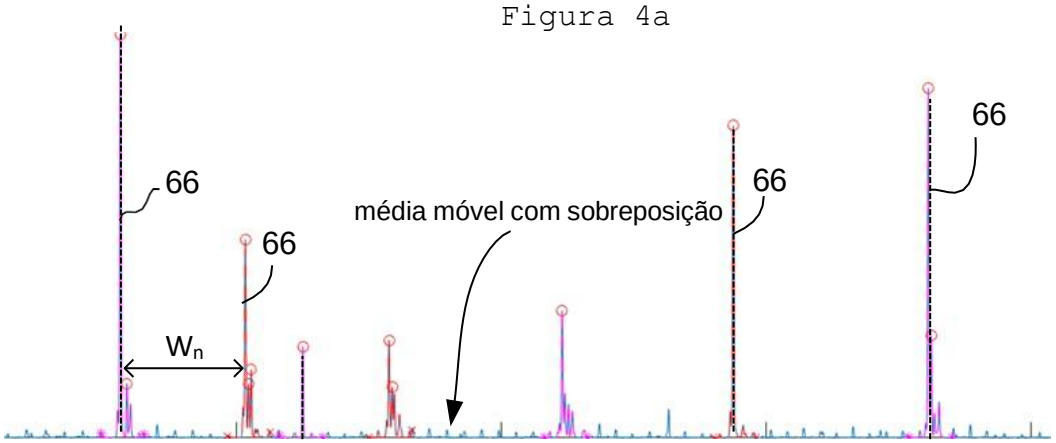


Figura 4b

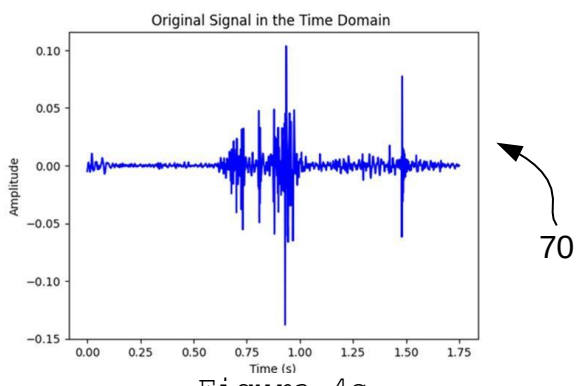


Figura 4c

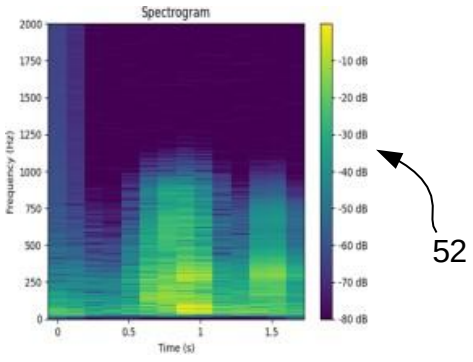


Figura 4d

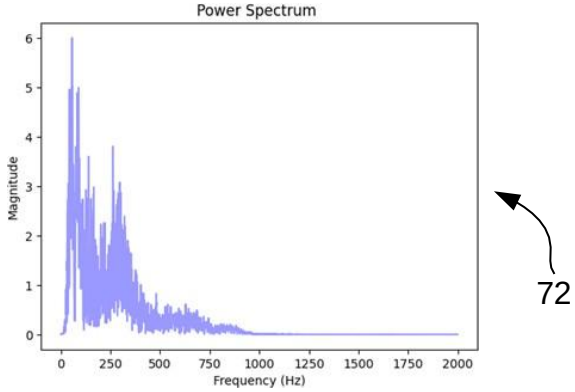


Figura 4e

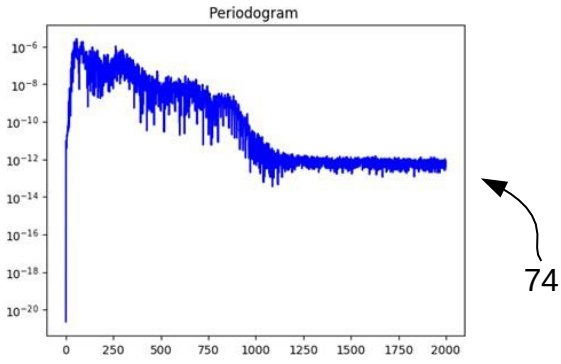


Figura 4f

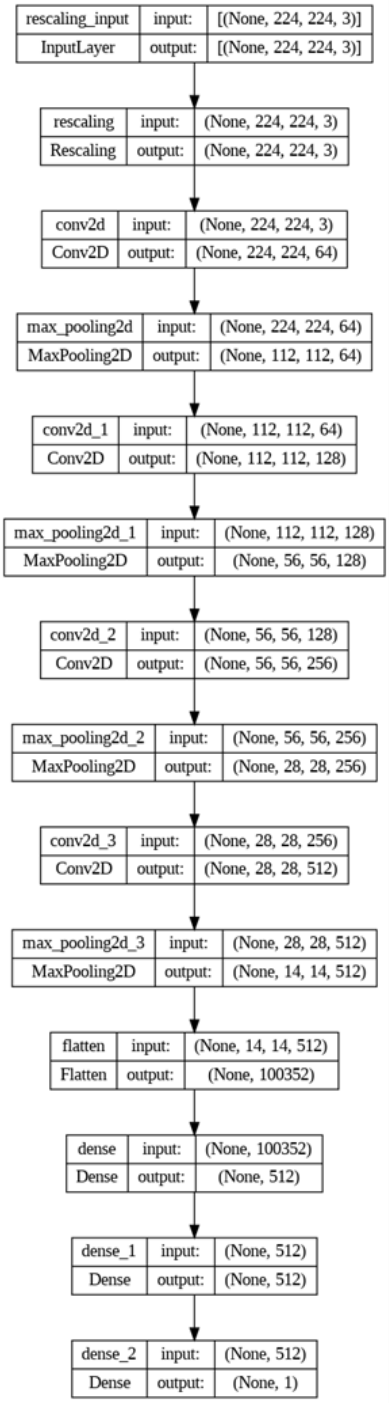


Figura 5a

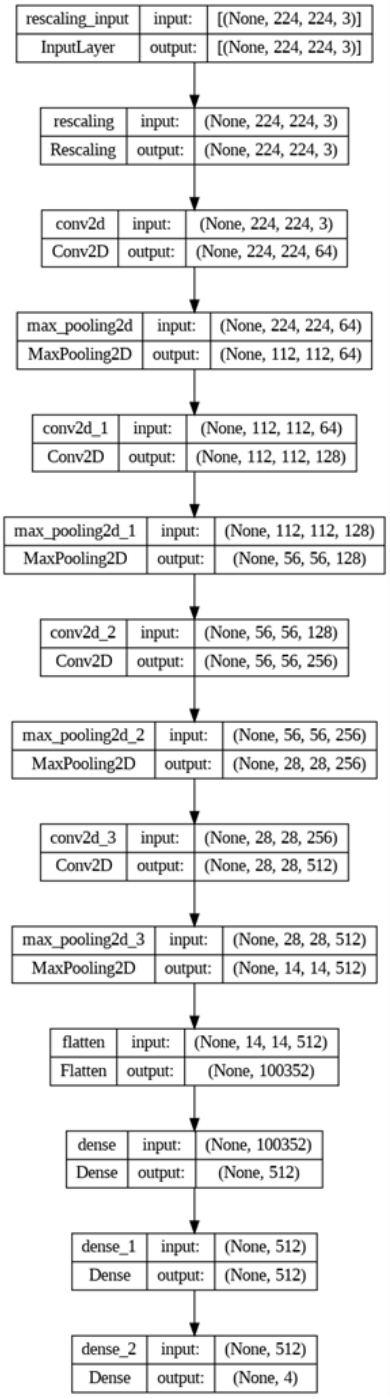


Figura 5b

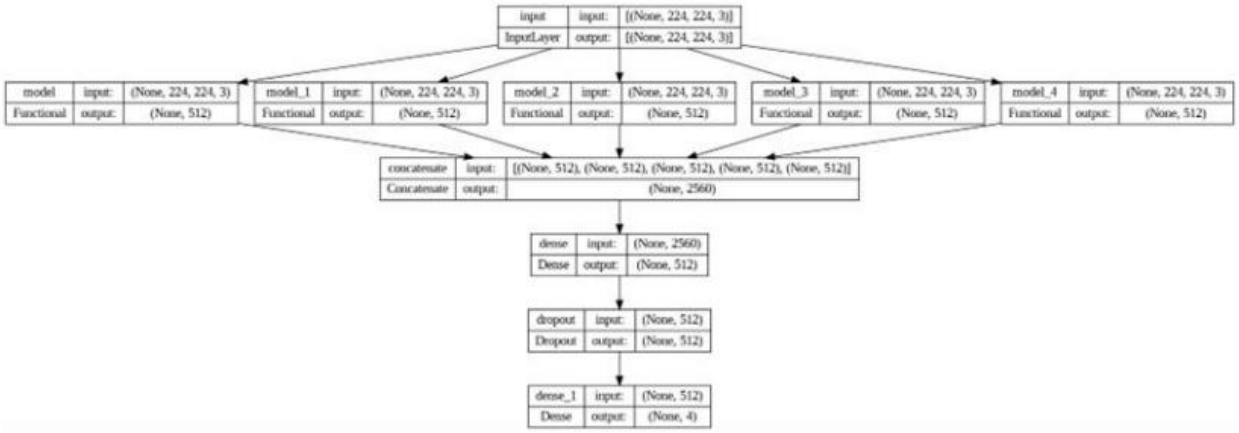


Figura 6

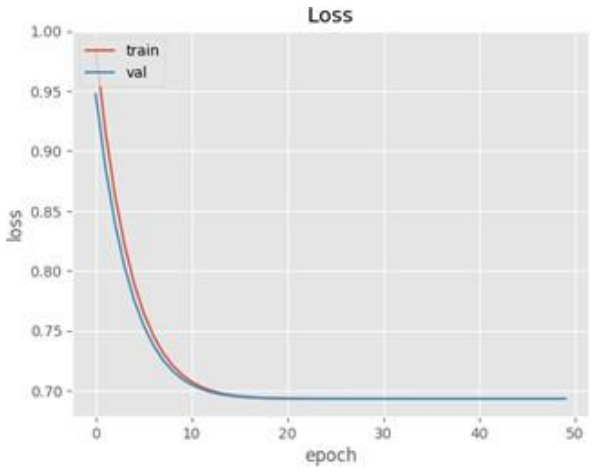


Figura 7a

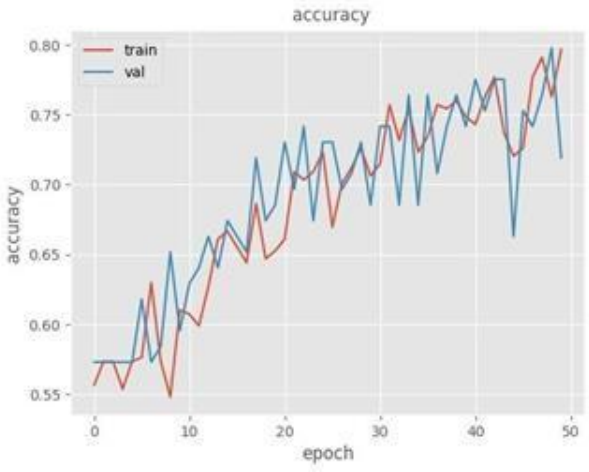


Figura 7b

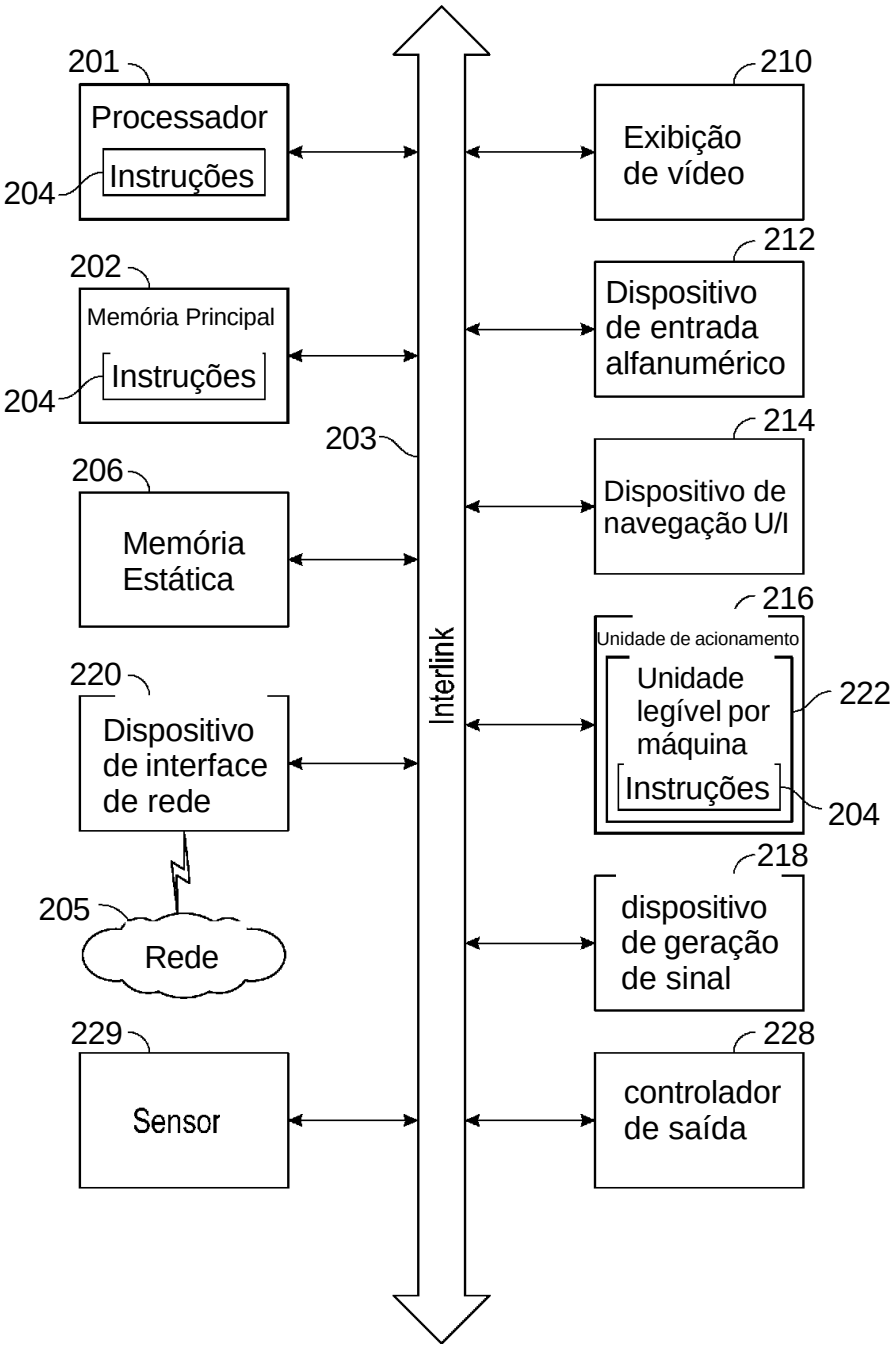

201
Processador
Instruções
204
202
Memória Principal
Instruções
204
206
Memória Estática
220
Dispositivo de interface de rede
205
Rede
229
Sensor
203
Interlink
210
Exibição de vídeo
212
Dispositivo de entrada alfanumérico
214
Dispositivo de navegação U/I
216
Unidade de acionamento
Unidade legível por máquina
Instruções
222
204
218
dispositivo de geração de sinal
228
controlador de saída
200

Figura 8

RESUMO

"MÉTODOS E SISTEMA PARA DIAGNÓSTICO DE DISFAGIA OROFARÍNGEA BASEADO EM INTELIGÊNCIA ARTIFICIAL"

Um sistema de suporte para o diagnóstico de disfagia orofaríngea, o sistema compreendendo um ou mais processadores acoplados a pelo menos um dispositivo de memória armazenando instruções executáveis por um ou mais processadores para pelo menos executar as operações de: com um dispositivo sensor, adquirir sinais acústicos emitidos durante um evento de deglutição de um material por uma pluralidade de indivíduos; identificar sinais acústicos de deglutição dentro dos sinais acústicos emitidos; pré-processar os sinais acústicos de deglutição; converter os sinais acústicos de deglutição em dados de imagem; extrair aspectos representativos do evento de deglutição dos dados de imagem; gerar um conjunto de dados compreendendo os dados de imagem; dividir os dados de imagem em um conjunto de dados de treinamento e um conjunto de dados de teste; inserir o conjunto de dados de treinamento em uma rede neural para treinar um modelo para classificar os dados de imagem com base em pelo menos um dos aspectos, como indicativo de deglutição normal e deglutição anormal devido à disfagia orofaríngea; validar o diagnóstico de disfagia orofaríngea usando dados de teste; e produzir um modelo treinado para pelo menos diagnosticar disfagia orofaríngea.