

Capítulo

6

Computação Visual e Detecção Precoce de Doenças em Escala Global: Oportunidades e Desafios

Manuel M. Oliveira¹, Giovani A. Meneguel¹, Maikel M. Rönnau¹,
Pantelis V. Rados²

¹ Programa de Pós-Graduação em Computação (PPGC) - UFRGS

² Programa de Pós-Graduação em Odontologia (PPGODO) - UFRGS

Abstract

Combining recent advances in machine learning with the Internet infrastructure and the computing capabilities of smartphones allows us to develop computational solutions with the potential to positively impact people's lives on a global scale. Such potential is particularly promising for the early detection of diseases, such as cancer. This work presents a software architecture for addressing these and other critical challenges in public health. It discusses how to develop solutions that benefit the population, leading to feedback to improve the quality of the provided services, thus creating a virtuous cycle.

Resumo

Avanços recentes em técnicas de aprendizagem de máquina combinados à infraestrutura da Internet e aos recursos disponíveis em smartphones nos permitem desenvolver soluções computacionais capazes de impactar positivamente a qualidade de vida das pessoas em escala planetária. Tal potencial é particularmente importante para a detecção precoce de doenças como, por exemplo, o câncer. Este trabalho apresenta uma arquitetura de software para o desenvolvimento de sistemas que visam enfrentar estes e outros importantes desafios em saúde pública. Ele discute como desenvolver soluções que beneficiem a população, que por sua vez provê realimentação para o aprimoramento dos serviços disponibilizados, criando um ciclo virtuoso.

Vídeo com a apresentação do capítulo: <https://youtu.be/114VMarSOsg>

6.1. Introdução

O uso de tecnologias da informação e comunicação (TICs) tem permitido ampliar o acesso da população a serviços de saúde, bem como racionalizar os custos associados à disponibilização desses serviços. Os termos *telemedicina* e *telessaúde* são utilizados para se referir a estas iniciativas que, segundo a Organização Mundial da Saúde (OMS), consistem na “*disponibilização de serviços de saúde, onde a distância é um fator crítico, e nos quais profissionais de saúde utilizam tecnologias da informação e comunicação para troca de informações visando o diagnóstico, tratamento e prevenção de doenças, pesquisa e avaliação, bem como para o processo de educação continuada de profissionais de saúde, objetivando a melhoria da saúde de indivíduos e de suas comunidades*” [WHO 1998]. Embora o termo telemedicina seja por vezes utilizado para referir-se especificamente a serviços prestados por médicos enquanto o termo telessaúde é utilizado para designar serviços prestados por profissionais de saúde em geral, neste trabalho os dois termos são usados como sinônimos.

Os primeiros registros de atividades de telemedicina remontam ao início do século 20 quando dados de eletrocardiograma foram transmitidos através de linhas telefônicas [Le 1906]. Atualmente, serviços como teleradiologia, teledermatologia, telepatologia, telepsiquiatria e mesmo telecirurgias estão se tornando comuns em países desenvolvidos [WHO 2010], e a recente pandemia do coronavírus (SARS-COV-2) contribuiu para uma rápida adoção de serviços de teleconsulta em diversos países. O Brasil, caracterizado por sua vastidão geográfica, por regiões de difícil acesso, e por uma distribuição irregular de serviços médicos especializados e de qualidade, possui um grande potencial para ampliação do acesso a serviços de saúde à sua população via telemedicina. Visando explorar este potencial, o Ministério da Saúde em articulação com as universidades públicas implementou núcleos técnico-científicos e criou o Programa Nacional Telessaúde Brasil Redes (PNTBR) [Maldonado et al. 2016].

Entretanto, a implementação de um amplo e efetivo programa de telessaúde requer a solução de diversos **desafios**, entre os quais:

1. Como garantir a disponibilização de serviços com ampla cobertura geográfica;
2. Como garantir a qualidade destes serviços e o seu constante aprimoramento;
3. Como garantir a escalabilidade dos serviços para um grande número de usuários;
4. Como prover serviços em diversas especialidades e para múltiplas doenças;
5. Como garantir a disponibilidade dos serviços 24 horas por dia, 7 dias por semana;
6. Como auxiliar a detecção precoce de doenças, visando a melhoria da qualidade de vida dos pacientes e a redução dos custos dos tratamentos;
7. Como implementar os requisitos acima com baixo custo.

Sem esquecer que os profissionais de saúde são os principais atores de um serviço de telessaúde, propomos uma estratégia para solução dos desafios técnicos listados acima. Tal

estratégia permitirá a disponibilização de sistemas de telessaúde mais efetivos, podendo ser diretamente incorporada ao PNTBR.

A busca por sistemas de saúde mais inclusivos, de qualidade, e com efetividade de custo guiou o desenvolvimento de sistemas de telessaúde na última década. Neste contexto, a Internet e os smartphones tiveram um papel fundamental. Recentemente, observamos o desenvolvimento acelerado de mais uma tecnologia disruptiva caracterizada por técnicas de aprendizagem de máquina, notadamente ligadas a áreas como visão computacional e modelos de linguagens. No caso específico de visão computacional, modelos baseados em redes neurais convolucionais (CNNs) têm demonstrado bons resultados na predição de diversas doenças como câncer de pele [Dildar et al. 2021], de mama [Mambou et al. 2018] e de pulmão [Bhatia et al. 2019], Covid-19 [Ozturk et al. 2020, Ismael and Şengür 2021], retinopatia diabética [Gargeya and Leng 2017], glaucoma [Abbas 2017], e Alzheimer [Ebrahimighahnavieh et al. 2020], para citar apenas alguns exemplos. O surgimento de técnicas automáticas com acurácia semelhante a de especialistas tem o potencial de transformar a telemedicina, ampliando significativamente a disponibilidade e o alcance dos serviços oferecidos. Tais avanços em técnicas de aprendizagem de máquina combinados à infraestrutura da Internet e aos recursos disponíveis em smartphones nos oferecem a **oportunidade de desenvolver soluções computacionais escaláveis para telessaúde capazes de impactar positivamente a qualidade de vida das pessoas em escala planetária**. Este potencial é particularmente importante na identificação precoce de doenças.

Este trabalho apresenta uma arquitetura de software para suporte à detecção precoce de doenças partir de imagens (Seção 6.2). Ele discute o desenvolvimento de soluções escaláveis, de baixo custo e com potencial de alcance mundial que beneficiem a população, que por sua vez provê realimentação para o aprimoramento dos serviços, criando um ciclo virtuoso. Para tal, a arquitetura proposta disponibiliza serviços automatizados de telessaúde (*e.g.*, predição de doenças) utilizando infraestrutura existente: a Internet, que fornece o meio de comunicação entre os provedores de solução e seus usuários; smartphones, que provêm os recursos computacionais para os usuários; e lojas de aplicativos (*e.g.*, *Google Play* e *App Store*), que oferecem canais de distribuição de apps (para acesso aos serviços) com capilaridade global. O uso desta infraestrutura permite que estes sistemas sejam disponibilizados com um baixo investimento financeiro. Para ilustrar o uso da arquitetura proposta, apresentaremos um sistema para suporte à detecção precoce de câncer de boca que se encontra em desenvolvimento no Programa de Pós-Graduação em Computação (PPGC) da UFRGS em parceria com a Faculdade de Odontologia (Seção 6.3).

6.2. Uma Arquitetura para Detecção Precoce de Doenças a partir de Imagens

A Figura 6.1 ilustra a arquitetura proposta, na qual profissionais de saúde podem submeter imagens de exames (*e.g.*, radiografias, retinografias, etc.) ou fotos de lesões suspeitas em seus pacientes (*e.g.*, manchas de pele, lesões na língua ou na mucosa bucal) para avaliação por um modelo de aprendizagem de máquina especializado no tipo de doença considerada (desafios 1, e 3 a 5). Tal modelo, treinado a partir de uma base de dados de imagens fornecidas e anotadas por especialistas, analisa cada imagem submetida e retorna sua predição sobre a probabilidade de ocorrência da doença específica, indicando as áreas suspeitas na imagem. Quando apropriado, o sistema recomenda ao profissional de saúde a realização

de biópsia para confirmação de casos suspeitos (*e.g.*, suspeita de câncer). No caso de realização de biópsia, espera-se que o profissional de saúde submeta o laudo ao sistema. De posse do laudo (positivo ou negativo) e do conjunto de imagens associadas fornecidas pelo profissional de saúde, um profissional responsável pela curadoria dos dados acrescentará as imagens correspondentes com as respectivas anotações ao banco de imagens. O banco de imagens então atualizado será utilizado para gerar (treinar) uma nova versão do modelo de predição. A disponibilização de novas imagens com seus respectivos laudos gera um ciclo virtuoso que deve levar a uma melhoria da acurácia do modelo de predição (desafio 2).

O sistema pode ser hospedado em um servidor institucional ou em algum serviço em nuvem. Completam este ecossistema os desenvolvedores de aplicativos (apps) para dispositivos móveis. Os apps são disponibilizados através das lojas de aplicativos, tornando-os acessíveis em todo o mundo (desafio 1). A interação do profissional de saúde com o modelo de predição é realizada por meio destes aplicativos ou por meio de uma aplicação web acessível via computadores pessoais (Figura 6.1). Utilizando estes mecanismos, os profissionais de saúde podem enviar imagens para predição, enviar laudos de biópsias, e acessar os resultados das predições para seus pacientes. As imagens enviadas são submetidas ao modelo de predição, sendo armazenadas juntamente com os resultados das avaliações nos registros dos respectivos pacientes no banco de dados.

A arquitetura mostrada na Figura 6.1 pode ser especializada para predição de diferentes tipos de doenças, diferindo apenas com relação ao modelo de predição e ao conjunto de imagens e respectivas anotações utilizadas para o treinamento do modelo (desafio 4). Assim, por exemplo, o modelo de inferência pode ser instanciado para detecção de doenças como câncer de pele [Dildar et al. 2021], câncer de mama [Mambou et al. 2018], câncer de pulmão [Bhatia et al. 2019], câncer de boca, para detecção de Covid-19 a partir de raio-X [Ozturk et al. 2020, Ismael and Şengür 2021], retinopatia diabética [Gargeya and Leng 2017], glaucoma a partir de imagens de fundo de olho [Abbas 2017], e Alzheimer a partir de neuroimagens [Ebrahimighahnavieh et al. 2020]. Por ser um serviço automatizado, disponível 24 horas por dia, apresentar um custo relativamente baixo, e ser aplicável a vários tipos de doenças, acreditamos que esta estratégia terá grande importância na democratização de serviços de saúde nos próximos anos. Note que um único servidor pode hospedar vários modelos, cada um treinado para detecção (precoce) de um tipo específico de doença. Isto contribui para uma redução ainda maior dos custos de disponibilização destes serviços (desafios 6 e 7).

Os smartphones constituem uma interface bastante conveniente de interação com esta arquitetura, permitindo, por exemplo, a captura e envio de imagens para avaliação pelo sistema, bem como o acompanhamento dos resultados. A Seção 6.3 descreve um sistema para detecção de câncer de boca baseado na arquitetura apresentada.

6.3. Sistema para Detecção Precoce de Câncer de Boca

O câncer é a segunda maior causa de mortes em todo o mundo [Roser and Ritchie 2015] e estima-se que em 2020 a doença tenha vitimado 10 milhões de pessoas [Sung et al. 2021]. O câncer de boca, por sua vez, é o tipo mais prevalente na região da cabeça e pescoço, com estimativas de 657.000 novos casos e 300.000 mortes anualmente, sendo

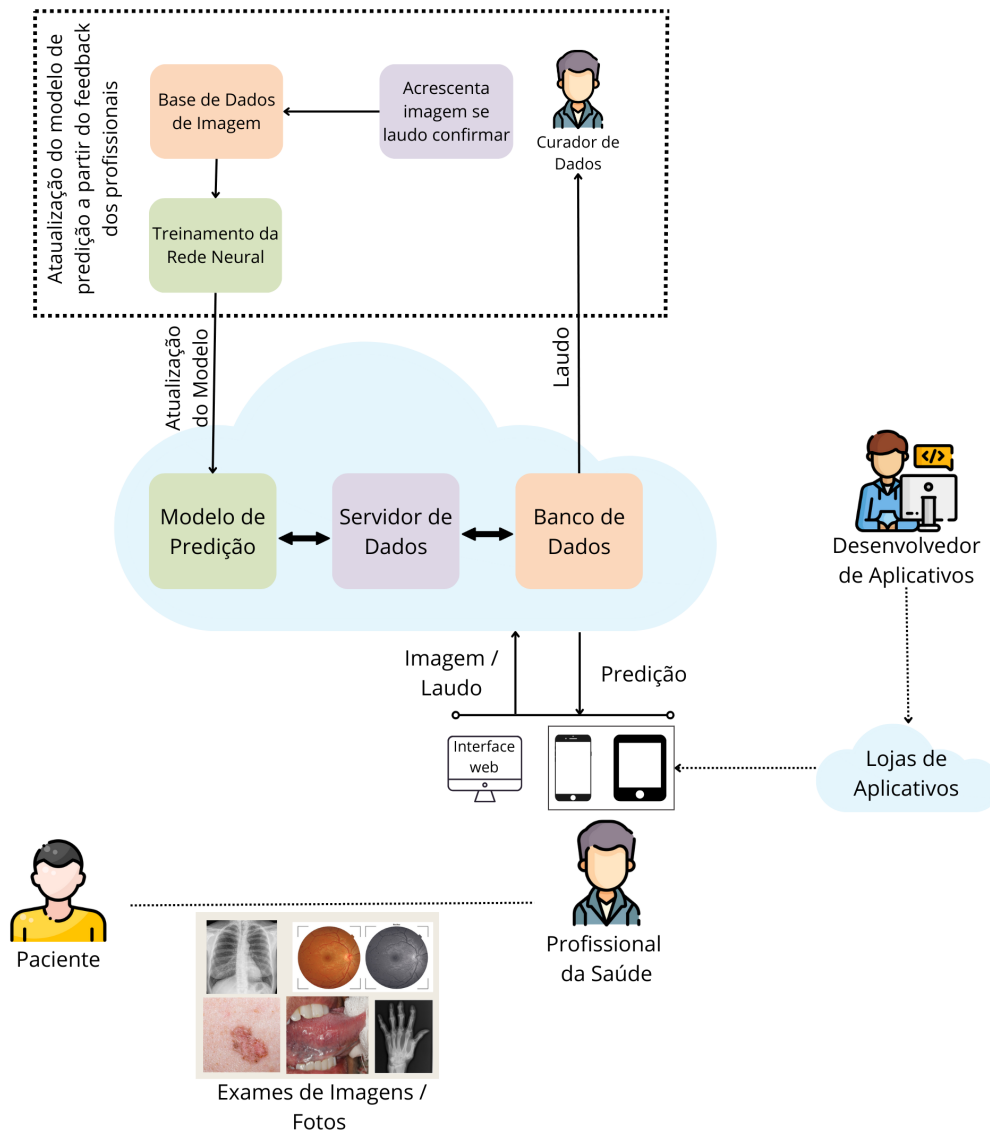


Figura 6.1: Arquitetura geral de sistema proposto para detecção precoce de doenças a partir de imagens.

que por mais de uma década o número de diagnósticos vem subindo [Foundation 2023]. Atualmente, a forma mais efetiva de reduzir o número de vítimas fatais de todas as formas de câncer é a detecção precoce.

O diagnóstico do câncer de boca ainda depende da análise clínica e de realização de biopsia. Geralmente, isso ocorre apenas em estágios avançados da doença, quando o paciente sofre com o desconforto causado pelo tumor e busca ajuda médica. Apesar dos recentes avanços em tratamentos e procedimentos cirúrgicos, a taxa de sobrevivência ainda é inferior a 60% em um período de cinco anos [Ries et al. 1998]. Além disso, devido ao diagnóstico tardio e à agressividade do tumor (e do tratamento), pacientes são frequentemente submetidos a remoção de tecidos que além de causar deformidades faciais, impactam a capacidade de falar, engolir e mastigar [Foundation 2023]. Isso resulta não apenas em sequelas cosméticas e funcionais, mas também deixa marcas emocionais.

Por outro lado, se diagnosticado precocemente, as chances de sobrevivência sobem para 80-90% [Foundation 2023].

Objetivando contribuir para a detecção precoce do câncer de boca, encontra-se em desenvolvimento no PPGC-UFRGS em parceria com a Faculdade de Odontologia um sistema (*Oral Cancer Screening – OCS*) baseado na arquitetura apresentada na Figura 6.1. O desenvolvimento deste projeto foi autorizado pelo Comitê de Ética em Pesquisa da UFRGS (parecer CAAE - 39212420.9.0000.5347). Um aplicativo para smartphones permite que dentistas, ao perceberem algo incomum durante o exame de um paciente, fotografem as regiões suspeitas e enviem as fotos para avaliação em nosso servidor (Figura 6.2). Conforme apresentado na Seção 6.2, os resultados das avaliações são então disponibilizados aos dentistas, indicando as regiões das fotografias que contém elementos suspeitos. Quando apropriado, a resposta do sistema também inclui uma recomendação para realização de biópsia. Fotos de lesões acompanhadas por laudos de biópsias são utilizadas para aprimorar o treinamento do modelo, contribuindo para melhorar sua acurácia.

A comunicação entre os componentes do sistema permite a integração de diferentes tecnologias que complementam suas funcionalidades. A solução é constituída por três módulos principais: *interface com usuário (front end)*, *servidor de dados (back end)* e *modelo de predição*. O sistema possibilita o armazenamento dos dados de usuários (dentistas), pacientes, imagens submetidas, resultados das predições, recomendações, e laudos submetidos, além de possuir uma interface para acessos a estes dados. A seguir, é apresentado um detalhamento dos três módulos principais, os quais encontram-se representados na Figura 6.3.

6.3.1. Interface com o Usuário

A interação dos usuários com o sistema ocorre prioritariamente via smartphone por meio de app distribuído através das lojas de aplicativos (veja Figura 6.1) e disponível para as plataformas Android e iOS. O sistema também pode ser acessado através de uma interface web por meio de uma aplicação desktop. O aplicativo móvel foi desenvolvido utilizando o *framework React Native* enquanto a aplicação desktop foi desenvolvida utilizando o *framework ReactJS*. O aplicativo móvel permite a captura de fotos de lesões de boca observadas pelo dentista utilizando a câmera do smartphone e o seu envio para avaliação pelo modelo de predição. Ele também oferece acesso a todas as demais funcionalidades do sistema, incluindo o envio de laudos referentes a exames e o acesso a dados de pacientes (*e.g.*, imagens e laudos submetidos e resultados de predições). Por se tratar de um sistema multiplataforma, todas essas funcionalidades, exceto a captura de imagens, também está disponível na versão desktop, a qual é disponibilizada para conveniência dos usuários. Embora o aplicativo seja disponibilizado livremente, o acesso aos serviços descritos requer cadastramento, o que pode ser feito via interface com o usuário. Um curador analisa as solicitações de cadastramento, verifica os dados e certifica o usuário (dentista), como forma de garantir que o sistema seja utilizado por profissionais habilitados. A partir da certificação, é possível fazer *login* e acessar todas as funcionalidades do sistema.

A Figura 6.2 ilustra, de modo simplificado, os passos envolvidos no processo de suporte à detecção precoce de câncer de boca. Neste exemplo, o aplicativo é utilizado pelo dentista para fotografar uma lesão observada na língua de um paciente e enviar a

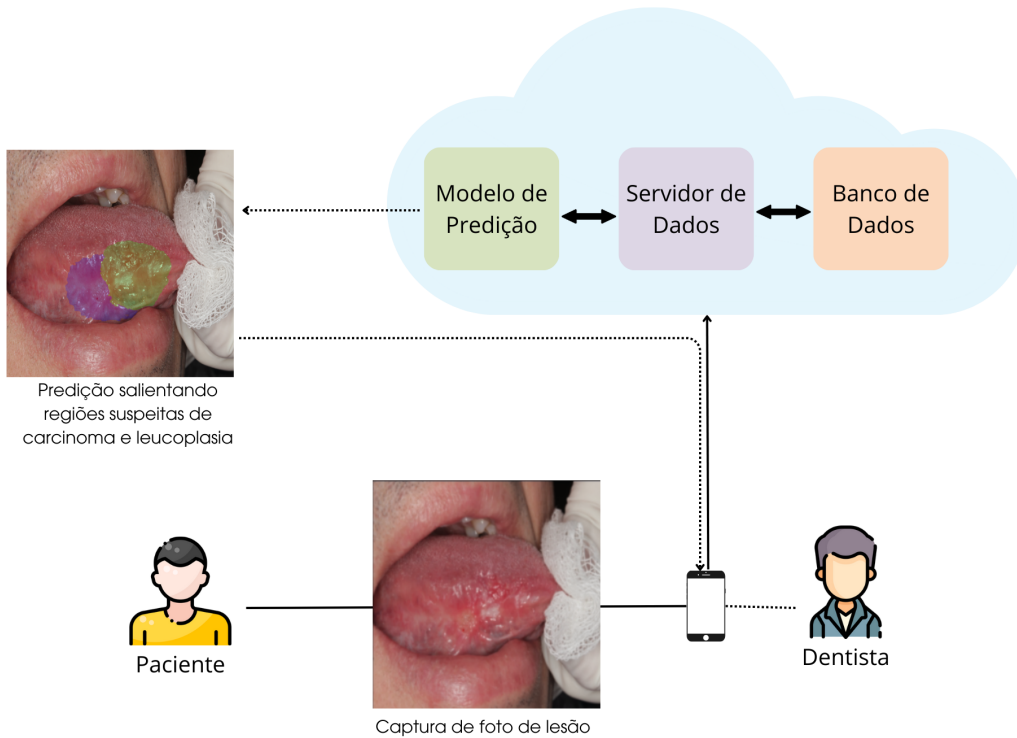


Figura 6.2: Visão simplificada do uso do sistema OCS. Dentista utiliza o smartphone para capturar foto de lesão na boca (língua) de paciente e enviá-la para avaliação pelo sistema. O resultado da predição é disponibilizado para o dentista, salientando regiões suspeitas.

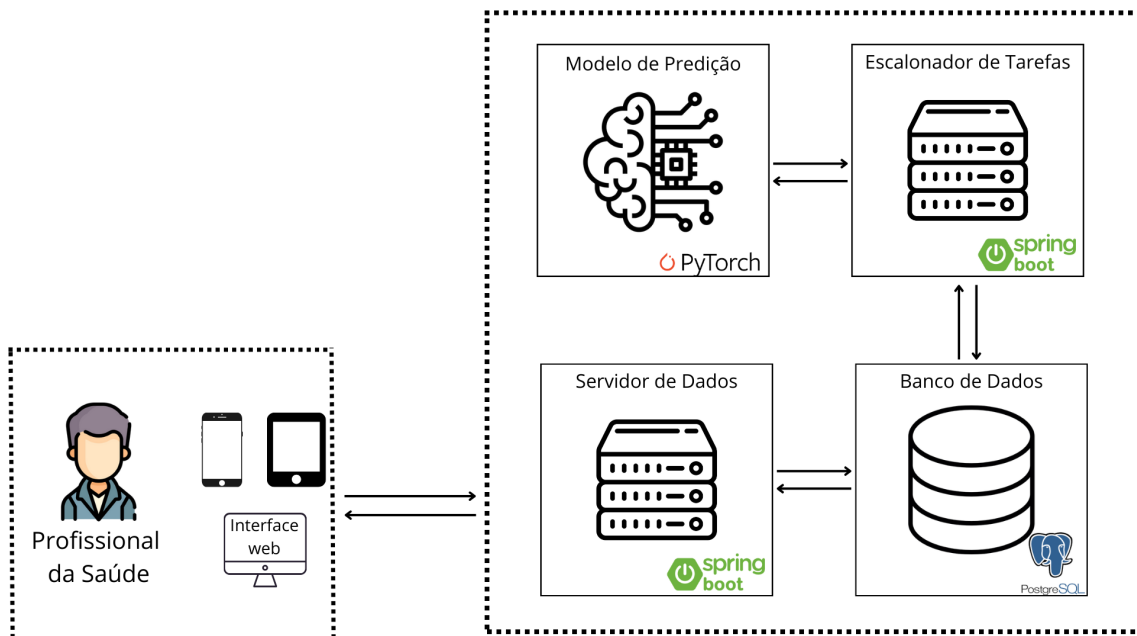


Figura 6.3: Componentes do Sistema OCS. Interface com o usuário (esquerda). Servidor de dados, modelo de predição, escalonador de tarefas e banco de dados (direita).

imagem para análise pelo sistema. O resultado da análise (predição) é disponibilizado para o dentista salientando regiões suspeitas de ocorrência de câncer e outras lesões ou

desordens potencialmente malignas de boca, que podem evoluir para câncer (Figura 6.2). Neste exemplo, o sistema recomenda ao dentista a realização de biópsia para confirmar as suspeitas. Espera-se que o dentista forneça o laudo da biópsia, para que um curador de dados possa acrescentar a imagem ao banco de imagens utilizado para treinamento do modelo (veja Figura 6.1), realimentando o processo com o objetivo de melhorar a acurácia do sistema. Note que tal retorno é importante independentemente do resultado da biópsia: em caso de existência de lesão maligna, a imagem reforçará a confiança do sistema na predição; caso contrário, a imagem contribuirá para ajustar a predição por meio de contraexemplo.

A Figura 6.4 apresenta algumas telas do aplicativo OCS. A Figura 6.4 (a) mostra a tela inicial (*login*). Após a verificação das credenciais (e-mail e senha), o usuário pode acessar os dados dos seus pacientes cadastrados (b), os quais são apresentados em ordem alfabética (c). A Figura 6.4 (d) mostra o resultado da predição realizada pelo modelo para uma foto de um dos pacientes.

Conforme mencionado na Seção 6.2, as lojas de aplicativos têm papel fundamental como canal de distribuição para tornar o sistema acessível a pessoas independente de suas localizações geográficas. Contribuindo para isso, o aplicativo e a aplicação web são disponibilizadas com versões em Português e em Inglês, podendo ser facilmente customizados para outros idiomas. Com o aumento do número de usuários, espera-se um aumento no número de realimentações (laudos) por parte dos usuários. Isto deverá contribuir para um incremento na quantidade e diversidade das imagens utilizadas para treinamento do modelo, com conseqüente impacto na acurácia do sistema.

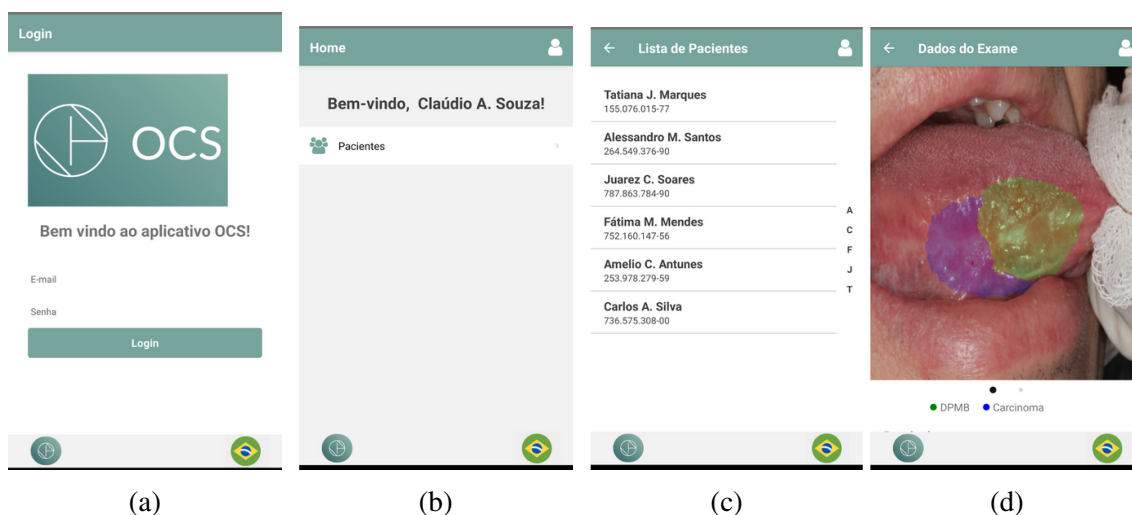


Figura 6.4: Exemplos de telas do aplicativo OCS para dispositivos móveis. (a) Tela inicial (*login*) solicitando dados de usuário e senha. (b) Tela mostrada após a verificação das credenciais do usuário. (c) Lista de pacientes (de um dentista). (d) Resultado da predição para uma foto de paciente, salientando regiões e indicando os tipos de lesões suspeitas.

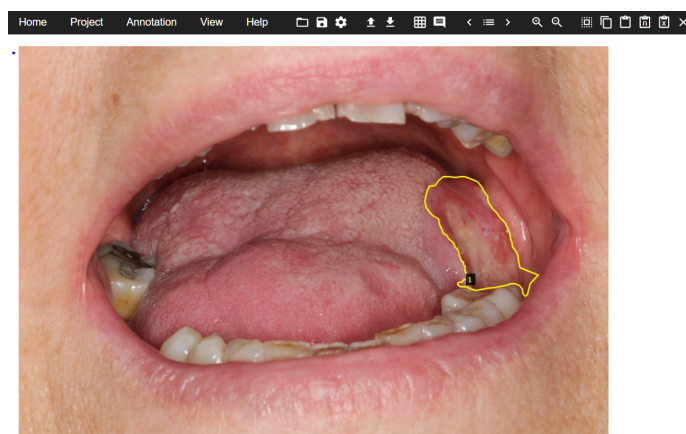


Figura 6.5: Exemplo de imagem anotada por especialista utilizando o software VIA [Dutta et al. 2016]. A região contendo uma lesão é delimitada por linha poligonal fechada simples e associada a um tipo de específico de lesão. Neste exemplo, trata-se de carcinoma espinocelular.

6.3.2. Servidor de Dados

Um banco de dados armazena as imagens, os resultados de predições e os laudos associados a cada paciente, os quais podem ser consultados pelos usuários através de um servidor de dados. Um escalonador de tarefas é responsável por verificar periodicamente a existência de solicitações de avaliação ainda não processadas e encaminhá-las, em ordem de chegada, para o modelo de predição (Figura 6.3). Os resultados das predições são coletados e armazenados pelo servidor no banco de dados.

O banco de dados foi implementado utilizando PostgreSQL, um gerenciador de bancos de dados relacionais disponibilizado como software livre e que apresenta características de robustez, segurança e extensibilidade. O servidor de dados foi desenvolvido utilizando o framework *Spring Boot* da linguagem Java. Este servidor possui um conjunto de funções (API) responsáveis por disponibilizar dados para os demais processos utilizando uma arquitetura cliente/servidor e protocolo de comunicação HTTP.

6.3.3. Modelo de Predição

O modelo de predição corresponde a uma rede neural convolucional treinada para detectar lesões de boca (Figura 6.4d). O treinamento está sendo realizado utilizando um conjunto de imagens (contendo lesões) fornecidas e anotadas por especialistas da Faculdade de Odontologia da UFRGS. Nestas imagens, cada lesão é delimitada por uma linha poligonal fechada e contém um código que identifica o tipo de lesão (Figura 6.5). A partir destas anotações, são geradas automaticamente máscaras que definem os pixels correspondentes a cada lesão. Estas máscaras são utilizadas juntamente com as imagens para o treinamento supervisionando do modelo. A versão atual do módulo de predição busca identificar duas classes de lesões: *carcinoma espinocelular* (câncer) e *leucoplasia* (uma forma de lesão que pode eventualmente evoluir para câncer). Em uma etapa posterior, pretende-se estender a classificação para incluir outras formas de desordens potencialmente malignas, neoplasias benignas, e outros tipos de lesões não suspeitas de malignidade.

6.3.4. Avaliação de Imagens de Microscopia

A arquitetura proposta também pode ser utilizada com imagens de microscopia para auxílio na detecção precoce de câncer de boca. Para tanto, basta realizar a substituição do modelo de predição.

A citopatologia pode ajudar a detectar os primeiros sinais de desenvolvimento de câncer de boca. O número de Regiões Organizadoras Nucleolares Argirófilas (AgNORs) encontradas no núcleo das células indica o quão rapidamente estas células estão se replicando e serve como um indicador de lesões com potencial maligno [Jajodia et al. 2017]. Dada seu menor custo em relação a outras técnicas, a coloração de AgNORs é uma opção atraente, especialmente para países em desenvolvimento. No entanto, a contagem manual, ainda utilizada atualmente, envolve o trabalho de um especialista (citopatologista), sendo um processo demorado, cansativo e sujeito a erros. Visando eliminar estas limitações, desenvolvemos um método automático que utiliza uma CNN para segmentar e contar o número de núcleos e de AgNORs em cada núcleo em imagens de lâminas de microscopia. O modelo resultante apresenta desempenho similar ao de citopatologistas, sendo, entretanto, significativamente mais rápido. A Figura 6.6 compara os resultados da segmentação automática de núcleos e AgNORs realizados pela técnica desenvolvida (Resultado) contra segmentações de referência (Referência) realizadas manualmente por especialistas para um conjunto de imagens de lâminas citológicas (Entrada). Os núcleos, AgNORs e o fundo são mostrados nas cores laranja, azul e cinza, respectivamente. Note o alto grau de concordância das segmentações. A variabilidade de cores, contraste, e nível de ruído nas imagens de entrada atesta a capacidade de generalização do modelo. Uma descrição detalhada desta técnica e de seus resultados podem ser encontradas em [Rönnau et al. 2023].

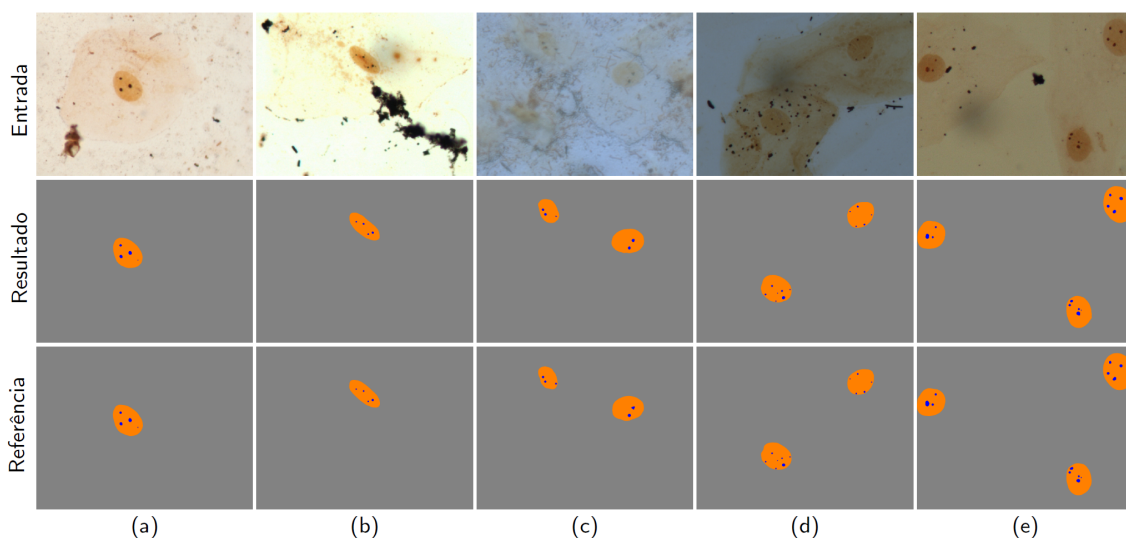


Figura 6.6: Exemplos de segmentação automática de núcleos e AgNORs em células da mucosa oral usando nosso modelo para AgNORs. Entrada: Imagens de lâminas citológicas coradas com AgNOR. Resultado: Segmentação automática produzida pelo nosso modelo. Referência: Segmentação manual realizada por especialistas (padrão ouro). Note o alto grau de concordância entre as segmentações.

Além da contagem de AgNORs, também investigamos a utilização de outras características celulares que possam, potencialmente, indicar a ocorrência de sinais precursoros de câncer de boca. A existência de núcleos com maior volume ou de aglomerados pode ser um indicador de malignidade. Para tentar detectar estes sinais, desenvolvemos mais um modelo preditivo baseado em CNN para classificação de núcleos e citoplasmas de células da mucosa bucal coradas pela técnica de Papanicolaou. A Figura 6.7 ilustra alguns resultados obtidos através de segmentação automática produzida pelo nosso modelo, comparando-os com a segmentação manual realizada por citopatologistas para um conjunto de imagens contendo características variadas. Observe-se, também neste caso, o alto grau de concordância entre as segmentações. Nos exemplos da Figura 6.7, a cor laranja corresponde ao citoplasma de uma célula isolada, ao passo que a cor azul (mais escuro) representa o citoplasma de aglomerados celulares (grupos de células em contato direto). Os núcleos representados em vermelho indicam células superficiais, ao passo que os núcleos em ciano estão associados a células intermediárias. O sistema também é capaz de detectar escamas (células anucleadas), bem como células suspeitas de malignidade (com núcleos com maior volume).

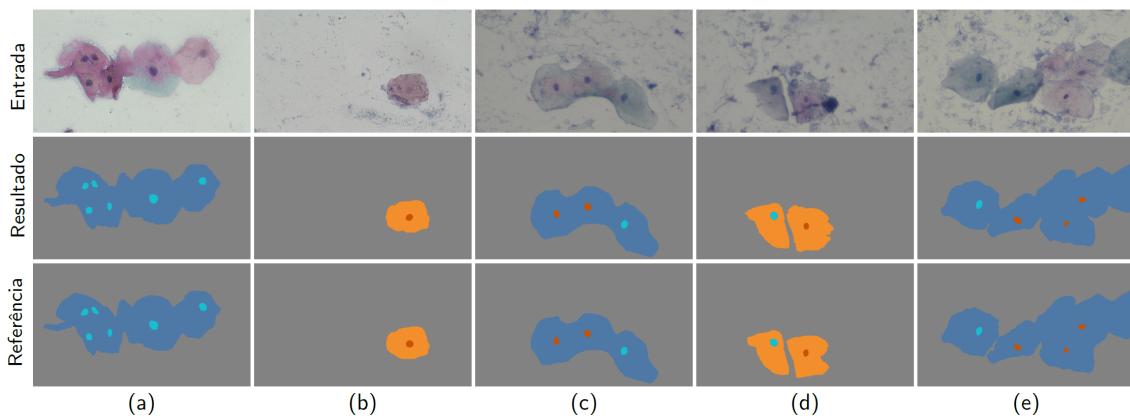


Figura 6.7: Exemplos de segmentação e classificação de células da mucosa bucal coradas pela técnica de Papanicolaou. Resultado: Segmentação e classificação automática produzida pelo nosso modelo. Referência: Segmentação e classificação manual produzidas por especialistas (padrão ouro). Mais uma vez, observe-se o elevado grau de concordância entre os dois resultados.

O processamento de imagens de células coradas pelas técnicas de AgNOR e Papanicolaou fornecem informações relevantes para a tomada de decisão de citopatologistas. A automatização desses processos permite a liberação dos profissionais para realização de outras tarefas. Também permite que os respectivos testes sejam realizados com maior rapidez e de maneira escalável, atendendo um maior número de pacientes.

6.4. Iniciativas Relacionadas

Esta seção discute algumas iniciativas relacionadas à arquitetura de sistema aqui descrito.

Haron et al. (2020) [Haron et al. 2020] desenvolveram um aplicativo para dispositivos móveis chamado de MeMoSA (*Mobile Mouth Screening Anywhere*) cujo objetivo é contribuir para a identificação de câncer de boca. O aplicativo permite a captura de fotos

da boca do paciente, as quais são enviadas para análise por especialistas que emitem um diagnóstico. O aplicativo não detecta regiões suspeitas. A análise de imagens é realizada por especialistas humanos, o que limita a escalabilidade desta solução. Além disso, os diagnósticos são emitidos a partir de análise de imagens sem a realização de biópsias.

Welikala et al. (2020) [Welikala et al. 2020] utilizaram a base de imagens capturadas com o aplicativo MeMoSA para treinar uma rede neural para detecção de câncer de boca. As imagens capturadas pelo aplicativo foram anotadas por especialistas e separadas em um conjunto de treinamento e um conjunto de teste. Não há registros de que a rede resultante tenha sido incorporada pelo projeto do MeMoSA ou disponibilizada como um serviço online.

Warin et al. (2021) [Warin et al. 2021] utilizaram uma CNN para classificação de existência ou não de câncer de boca. A CNN foi treinada com 700 imagens anotadas por especialistas e divididas igualmente entre imagens com e sem câncer de boca. Em um trabalho subsequente dos mesmos autores [Warin et al. 2022], as imagens são classificadas entre contendo desordens potencialmente malignas, contendo carcinoma, ou sem identificação de patologias. Huang et al. (2023) [Huang et al. 2023] treinaram uma CNN para classificação de imagens com relação à presença de câncer de boca utilizando 130 imagens disponíveis na plataforma Kaggle [Dataset 2023]. Lin et al. (2021) [Lin et al. 2021] utilizaram uma CNN (HRNet [Wu et al. 2021]), treinada a partir de imagens capturadas com diferentes modelos de smartphones e anotadas por especialistas, para classificar lesões de boca. A CNN classifica as imagens entre contendo: carcinoma, úlcera aftosa, mucosa normal, e lesões potencialmente malignas de alto e de baixo risco. Todos estes trabalhos descrevem o treinamento de CNNs que foram avaliadas de modo isolado utilizando apenas seus respectivos conjuntos de testes. Não há registros de que essas redes tenham sido utilizadas em serviços online.

Considerando outras áreas em telessaúde, Hacısofıtaoglu et al. (2020) [Hacısofıtaoglu et al. 2020] utilizaram uma CNN para a identificação de retinopatia diabética em imagens de fundo de olho. As imagens para treinamento foram capturadas usando diferentes modelos de câmeras do tipo *digital single lens reflex* (DSLR). Porém, o objetivo é fornecer como entrada para a CNN imagens capturadas por smartphones, o que é feito com o auxílio de diferentes equipamentos oftalmológicos. A classificação feita pela CNN é binária, ou seja, a imagem é classificada como apresentando ou não retinopatia diabética.

Archibong et al. (2017) [Archibong et al. 2017] utilizaram dispositivos móveis e técnicas convencionais de processamento de imagens (*i.e.*, sem o uso de aprendizagem de máquina) para a identificação de hemólise (processo de dissolução ou destruição de glóbulos vermelhos do sangue). Este processo é caracterizado pela elevada presença de enzimas hepáticas e baixa contagem de plaquetas. As imagens foram capturadas por um smartphone acoplado a um dispositivo que contém a amostra e processadas pelo smartphone. O resultado é exibido na tela do aparelho. O processo requer calibração da câmera, que é feita através de uma curva de calibração específica para cada modelo de smartphone.

Várias das iniciativas descritas acima utilizam CNNs para classificação de imagens e algumas utilizam smartphones para captura de imagens. Entretanto, nenhuma de-

las fornece uma solução integrada e escalável para detecção automática de doenças como mostrado na Figura 6.1.

6.5. Conclusão

Este trabalho apresentou uma proposta de arquitetura para construção de sistemas de tele-saúde escaláveis que podem ser customizados para detecção de diversos tipos de doenças. A solução proposta combina três importantes componentes: (i) modelos de aprendizagem de máquina, os quais podem ser treinados e customizados para identificar uma grande variedade de doenças; (ii) a Internet, que disponibiliza a infraestrutura de comunicação; e (iii) os smartphones e aplicativos, que oferecem uma interface versátil entre os usuários e os serviços disponibilizados. Esta combinação, ilustrada na Figura 6.1, permite a oferta de serviços para suporte à detecção de diversos tipos de doenças de forma automatizada, ininterrupta, e com baixo custo.

A arquitetura proposta se beneficia da realimentação provida por seus usuários, permitindo que o sistema evolua, melhorando assim a acurácia de suas predições. Sua implementação envolve custos relativamente modestos. Estes incluem: (i) a disponibilização de um servidor para hospedagem do banco de dados, do servidor de dados e do modelo de predição; (ii) a construção de um ou mais bancos de imagens anotadas por especialistas para doenças específicas; (iii) a especificação e treinamento de um ou mais modelos de predição a partir dos dados anotados; e (iv) o desenvolvimento do servidor de dados, de um ou mais aplicativos para dispositivos móveis, e, opcionalmente, de uma aplicação web para desktop.

Um sistema para suporte à detecção precoce de câncer de boca baseado na arquitetura proposta está sendo desenvolvido no PPGC em parceria com a Faculdade de Odontologia da UFRGS. Os componentes desenvolvidos para este sistema (aplicativo para dispositivo móvel, aplicação web para desktop, servidor de dados, e CNN para predição) podem ser utilizados como referência para o desenvolvimento de novos sistemas.

A possibilidade de treinar modelos de aprendizagem de máquina para detecção precoce de doenças como o câncer pode contribuir para evitar a ocorrência de mortes prematuras e para uma melhoria da qualidade de vida de pacientes. A estratégia aqui descrita representa uma evolução natural e necessária aos serviços de tele-saúde. Dada a sua aplicabilidade a vários cenários, baixo custo, facilidade de implementação e potencial de alcance em escala global, acreditamos que ela pode desempenhar um papel relevante na ampliação e democratização de serviços de saúde nos próximos anos.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. O presente trabalho foi realizado com apoio do CNPq [305474/2022-7], FAPERGS [51956.616.19834].

Referências

[Abbas 2017] Abbas, Q. (2017). Glaucoma-deep: detection of glaucoma eye disease on retinal fundus images using deep learning. *International Journal of Advanced Compu-*

- ter Science and Applications, 8(6).
- [Archibong et al. 2017] Archibong, E., Konnaiyan, K. R., Kaplan, H., and Pyayt, A. (2017). A mobile phone-based approach to detection of hemolysis. *Biosensors and bioelectronics*, 88:204–209.
- [Bhatia et al. 2019] Bhatia, S., Sinha, Y., and Goel, L. (2019). Lung cancer detection: a deep learning approach. In *Soft Computing for Problem Solving: SocProS 2017, Volume 2*, pages 699–705. Springer.
- [Dataset 2023] Dataset, K. (2023). Mouth cancer images. <https://www.kaggle.com/datasets/edward112/mouth-cancer-images>. Setembro de 2023.
- [Dildar et al. 2021] Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., Alsaiari, S. A., Saeed, A. H. M., Alraddadi, M. O., and Mahnashi, M. H. (2021). Skin cancer detection: a review using deep learning techniques. *International journal of environmental research and public health*, 18(10):5479.
- [Dutta et al. 2016] Dutta, A., Gupta, A., and Zissermann, A. (2016). VGG image annotator (VIA). <http://www.robots.ox.ac.uk/vgg/software/via/>. Version: 2.0.12, Accessed: 2023.
- [Ebrahimighahnavieh et al. 2020] Ebrahimighahnavieh, M. A., Luo, S., and Chiong, R. (2020). Deep learning to detect alzheimer’s disease from neuroimaging: A systematic literature review. *Computer methods and programs in biomedicine*, 187:105242.
- [Foundation 2023] Foundation, T. O. C. (2023). Cancer screening protocols. <https://oralcancerfoundation.org/discovery-diagnosis/cancer-screening-protocols/>. Julho de 2023.
- [Gargeya and Leng 2017] Gargeya, R. and Leng, T. (2017). Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*, 124(7):962–969.
- [Hacisoftoglu et al. 2020] Hacisoftoglu, R. E., Karakaya, M., and Sallam, A. B. (2020). Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems. *Pattern recognition letters*, 135:409–417.
- [Haron et al. 2020] Haron, N., Zain, R. B., Ramanathan, A., Abraham, M. T., Liew, C. S., Ng, K. G., Cheng, L. C., Husin, R. B., Chong, S. M. Y., Thangavalu, L. A., et al. (2020). m-health for early detection of oral cancer in low-and middle-income countries. *Telemedicine and e-Health*, 26(3):278–285.
- [Huang et al. 2023] Huang, Q., Ding, H., and Razmjoooy, N. (2023). Optimal deep learning neural network using issa for diagnosing the oral cancer. *Biomedical Signal Processing and Control*, 84:104749.
- [Ismael and Şengür 2021] Ismael, A. M. and Şengür, A. (2021). Deep learning approaches for covid-19 detection based on chest x-ray images. *Expert Systems with Applications*, 164:114054.

- [Jajodia et al. 2017] Jajodia, E., Raphael, V., Shunyu, N. B., Ralte, S., Pala, S., and Jitani, A. K. (2017). Brush cytology and agnor in the diagnosis of oral squamous cell carcinoma. *Acta cytologica*, 61(1):62–70.
- [Le 1906] Le, E. (1906). Le télécardiogramme [the telecardiogram]. *Arch. Int. Physiol.*, 4:132–164.
- [Lin et al. 2021] Lin, H., Chen, H., Weng, L., Shao, J., and Lin, J. (2021). Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis. *Journal of Biomedical Optics*, 26(8):086007–086007.
- [Maldonado et al. 2016] Maldonado, J. M. S. d. V., Marques, A. B., and Cruz, A. (2016). Telemedicine: challenges to dissemination in brazil. *Cadernos de saude publica*, 32:e00155615.
- [Mambou et al. 2018] Mambou, S. J., Maresova, P., Krejcar, O., Selamat, A., and Kuca, K. (2018). Breast cancer detection using infrared thermal imaging and a deep learning model. *Sensors*, 18(9):2799.
- [Ozturk et al. 2020] Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., and Rajendra Acharya, U. (2020). Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in Biology and Medicine*, 121:103792.
- [Ries et al. 1998] Ries, L. A. G., Kosary, C., Hankey, B., Miller, B., and Edwards, B. (1998). Seer cancer statistics review, 1973-1995. *Bethesda, MD: National Cancer Institute*.
- [Rönnau et al. 2023] Rönnau, M. M., Lepper, T. W., Amaral, L. N., Rados, P. V., and Oliveira, M. M. (2023). A cnn-based approach for joint segmentation and quantification of nuclei and nors in agnor-stained images. *Computer Methods and Programs in Biomedicine*, 242:107788.
- [Roser and Ritchie 2015] Roser, M. and Ritchie, H. (2015). Cancer. *Our World in Data*. <https://ourworldindata.org/cancer>.
- [Sung et al. 2021] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3):209–249.
- [Warin et al. 2021] Warin, K., Limprasert, W., Suebnukarn, S., Jinaporntham, S., and Jantana, P. (2021). Automatic classification and detection of oral cancer in photographic images using deep learning algorithms. *Journal of Oral Pathology & Medicine*, 50(9):911–918.
- [Warin et al. 2022] Warin, K., Limprasert, W., Suebnukarn, S., Jinaporntham, S., Jantana, P., and Vicharueang, S. (2022). Ai-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer. *Plos one*, 17(8):e0273508.

- [Welikala et al. 2020] Welikala, R. A., Remagnino, P., Lim, J. H., Chan, C. S., Rajendran, S., Kallarakkal, T. G., Zain, R. B., Jayasinghe, R. D., Rimal, J., Kerr, A. R., et al. (2020). Automated detection and classification of oral lesions using deep learning for early detection of oral cancer. *IEEE Access*, 8:132677–132693.
- [WHO 1998] WHO (1998). *A Health Telematics Policy in Support of WHO'S Health-For-All Strategy for Global Development: Report of the WHO Group Consultation on Health Telematics*. World Health Organization.
- [WHO 2010] WHO (2010). *Telemedicine: opportunities and developments in member states. Report on the second global survey on eHealth*. World Health Organization.
- [Wu et al. 2021] Wu, H., Liang, C., Liu, M., and Wen, Z. (2021). Optimized hrnet for image semantic segmentation. *Expert Systems with Applications*, 174:114532.