

Protótipo de Aplicação Mobile Para Identificação de Úlceras do Pé Diabético

Rodrigo N. Borges¹, Rodrigo M. S. Veras¹, Elineide Santos¹

¹Departamento de Computação,
Universidade Federal do Piauí - Teresina, Brasil

{r.borges, rveras}@ufpi.edu.br, elineide.silva.inf@gmail.com

Abstract. *This article discusses the development of a mobile application prototype for healthcare professionals that uses Convolutional Neural Networks to detect lesions in patients with Diabetic Foot Ulcers. The article describes the development methodology, presents the prototype and the CNN model. The YOLOv5 network is used for lesion detection, and the results show that the model exhibited a Mean Average Precision of 86.20%.*

Resumo. *Este artigo discute o desenvolvimento de um protótipo de aplicação móvel para profissionais de saúde que utiliza Redes Neurais Convolucionais para detectar lesões em pacientes com Úlceras do Pé Diabético. O artigo descreve a metodologia de desenvolvimento, apresenta o protótipo e o modelo CNN. A rede YOLOv5 é utilizada para detecção das lesões, e os resultados mostram que o modelo exibiu um Mean Average Precision de 86,20%.*

1. Introdução

O Diabetes Mellitus é uma doença crônica que afeta milhões de pessoas em todo o mundo [Saeedi et al. 2019], e uma de suas complicações são as úlceras causadas pela dificuldade de cicatrização presente em indivíduos. Um dos principais tipos é conhecido como Úlcera do Pé Diabético (DFU). DFUs podem levar a sérias consequências como hospitalização, amputação e até morte, tornando-os um problema de saúde significativo.

Para melhorar a avaliação das lesões, foram propostas soluções baseadas no treinamento de Redes Neurais Convolucionais (CNNs) para reconhecer diferentes tipos de lesões, como infecção e isquemia, além de rastrear a progressão/regressão de úlceras [Goyal et al. 2020]. A combinação dessas redes pode trazer melhores resultados [Galdran et al. 2021], oferecendo subsídios valiosos ao diagnóstico de DFU.

Tendo isso em vista, este projeto visa elaborar o protótipo de um software móvel voltado para profissionais da saúde que aplique as soluções baseadas em CNN e possibilite a detecção das lesões. Espera-se que o trabalho seja a primeira etapa na construção de uma aplicação robusta que otimize as prescrições de medicamentos, melhore o monitoramento de lesões e melhore a segurança geral do tratamento do paciente.

A Seção 2 fornece uma revisão da literatura existente sobre o assunto. A Seção 3 descreve a metodologia, incluindo a arquitetura inicial, ferramentas utilizadas e aplicação de modelos de reconhecimento. As seções 4 e 5 apresentam resultados parciais alcançados e por fim, a Seção 6 discute a evolução do projeto e perspectivas futuras.

2. Revisão da Literatura

[Galdran et al. 2021] apresenta uma comparação do desempenho de CNNs e Vision Transformers no contexto do desafio DFUC 2021. É destacado o desempenho superior das CNNs em um regime de poucos dados, o que é atribuído à sua capacidade de explorar melhor as correlações espaciais. Também é demonstrada a eficácia do algoritmo Sharpness-Aware Minimization (SAM) em melhorar a capacidade de generalização de ambos os modelos.

A abordagem utilizada por [Santos et al. 2022] para classificar imagens de DFUs combinou cinco redes neurais convolucionais modificadas - VGG-16, VGG-19, Resnet-50, InceptionV3 e Densenet-201. As CNNs foram ajustadas e foram avaliadas configurações de camadas totalmente conectadas, além de normalização de lote e operações de dropout para definir os parâmetros. Os resultados mostraram que as CNNs modificadas se adequaram bem ao problema, mas a união das cinco CNNs aumentou significativamente as taxas de sucesso.

[Ploderer et al. 2018] é um estudo que desenvolveu uma aplicação mobile para envolver as pessoas com diabetes no autocuidado. O aplicativo apresenta metas, monitoramento de progresso e lembretes, bem como análises visuais que extraem e monitoram informações de tamanho de úlcera de fotos de celular do pé. O software foi avaliado por meio de um processo de design centrado no usuário e os resultados mostraram recursos promissores para envolver as pessoas no autocuidado.

A partir das obras analisadas fica evidente que o uso de Redes Neurais Convolucionais é uma alternativa benéfica para solucionar este tipo de problema. Também é possível verificar a existência de algumas aplicações mobile que têm objetivos semelhantes à que está sendo proposta neste projeto e que precisam de evolução em determinados aspectos. Sendo assim confirma-se a lacuna que pode ser preenchida pela aplicação proposta.

3. Metodologia de desenvolvimento do software

Nesta seção são exibidos os resultados da elaboração do projeto. O processo até então envolveu um estudo intensivo da literatura relacionada a fim de esboçar um esquema sólido e completo no qual a aplicação pudesse ser desenvolvida. A seguir serão apresentados os resultados da definição de requisitos, escolha de padrões de projeto, arquitetura do software e escolha de linguagens e frameworks.

3.1. Requisitos

Dentre todos os requisitos levantados, destacam-se principalmente os seguintes requisitos funcionais:

1. O software deve ser compatível com todos os principais sistemas operacionais móveis (Android e iOS) e acessível em vários tamanhos de tela de dispositivo.
2. O aplicativo deve possuir um recurso de câmera que permita aos profissionais de saúde capturar imagens da úlcera e áreas adjacentes a serem analisadas
3. A aplicação deve usar os modelos de reconhecimento baseados em CNNs para analisar as imagens, fornecendo medições precisas da evolução da lesão e categorização da úlcera com base na gravidade, de acordo com os rótulos estabelecidos no modelo.

Destacam-se também os seguintes requisitos não-funcionais:

1. O aplicativo deve ter um tempo de resposta rápido e ser capaz de lidar com grandes quantidades de dados sem ficar lento ou travar.
2. A aplicação deve ser confiável e estar sempre disponível, com o mínimo de tempo de inatividade ou erros.

3.2. Padrões de Projeto e Arquitetura

O padrão Model-View-Controller (MVC) é um padrão de design amplamente utilizado para o desenvolvimento de aplicações centradas no usuário [Pop and Samuel 2014]. É particularmente útil para aplicativos que exigem uma separação de preocupações entre a interface do usuário, a lógica de negócios e os dados.

Tendo em vista as necessidades presentes no presente projeto, o padrão MVC mostra-se uma escolha adequada devido aos fatores descritos anteriormente. Será, portanto, o padrão de projeto escolhido para a construção da aplicação.

A arquitetura proposta para o projeto tem como base a escolha do padrão de projeto MVC descrito na subseção anterior. Tendo isso em vista é possível partir do ponto inicial dos três componentes principais e, a partir disso, desenvolver um sistema mais robusto.

A camada de interface do usuário, referente à View, a camada Model deve lidar com o armazenamento e a recuperação dos dados do paciente relacionados às lesões, a camada Controller deve gerenciar a lógica de negócios do aplicativo.

Além destes três componentes, foi decidido adicionar algumas camadas que serão úteis para auxiliar o funcionamento do projeto. A camada denominada por Serviços deve lidar com as interações com os serviços externos exigidos pelo aplicativo, a camada de persistência deve fornecer certa abstração sobre o armazenamento de dados feito pelo aplicativo e finalmente, a camada de utilitários deve fornecer um conjunto de classes e componentes que podem ser usados em toda a aplicação.

3.3. Linguagens e Frameworks

A fim de se atender aos requisitos levantados é necessário fazer uso de tecnologias e ferramentas com recursos adequados ao problema. Nesse contexto, tem-se a escolha do uso do framework React Native que utiliza a linguagem JavaScript, bem como da base de dados Firebase.

Desse modo, o React Native [Facebook 2023] e o Firebase [Google 2023], juntamente com o modelo CNN para classificação de imagens, podem ser projetados para seguir o padrão de design MVC, no qual o React Native serve como a camada View, o Firebase serve como a camada Model e as APIs que interagem com o modelo de reconhecimento servem como uma das partes da camada Controller. Essa abordagem pode fornecer uma clara separação de componentes e tornar o aplicativo mais modular e escalável.

4. Detecção das lesões

Esta seção descreve a metodologia aplicada à detecção de lesões, abrangendo tópicos que detalham o conjunto de dados usado, a arquitetura CNN utilizada e as métricas de avaliação selecionadas para avaliá-la.

4.1. Base de Imagens

O conjunto de dados DFUC2020, apresentado por [Yap et al. 2020] é heterogêneo, incluindo variações de distância, ângulo, orientação, iluminação, foco e presença de objetos de fundo. As imagens possuem dimensões que variam de 1600×1200 a 3648×2736 pixels. Para encontrar um equilíbrio entre a eficiência do treinamento e a qualidade da imagem, todas as imagens foram reduzidas uniformemente para 640×480 pixels.

Dentro do conjunto de treinamento, há um total de 2.496 ocorrências anotadas de DFUs. Para maior utilidade, o conjunto de treinamento é acompanhado por um arquivo CSV contendo as coordenadas x e y (máximo e mínimo) que definem as caixas delimitadoras para cada úlcera. A versão publicada em 2020 deste conjunto - que foi utilizada neste projeto - possui imagens de classe única.

O conjunto de dados consiste em 4.000 imagens, que são divididas em dois conjuntos distintos:

- DFUC2020 train: Este conjunto contém 2.000 imagens que retratam exclusivamente úlceras de pé diabético.
- DFUC2020 test: Composto por outras 2.000 imagens, este conjunto abrange lesões e diversas outras condições dos pés, incluindo imagens de pés saudáveis.

4.2. Algoritmo para Detecção

A rede YOLOv5 [Jocher et al. 2022], é uma implementação da estrutura You Only Look Once [Redmon et al. 2016] que apresenta uma composição arquitetônica baseada em três componentes principais: *backbone*, *neck* e *head*. O componente *backbone* captura features multi escala a fim de reconhecer objetos de dimensões variadas.

O módulo *neck* apresenta refinamentos que reforçam a precisão da localização do objeto e a exatidão da classificação. Esse aumento contribui para um processo de detecção mais robusto e confiável.

Por sua vez, o componente *head* é responsável pela parte de predição na YOLOv5. Ele designa bounding boxes, classes de objetos e pontuações de confiança, aproveitando as caixas de âncora para explicar a diversidade de escalas e proporções de objetos. Essa abordagem garante a adaptabilidade do modelo a um amplo espectro de possíveis cenários de detecção.

Uma característica essencial deste modelo é sua eficiência obtida por meio de uma arquitetura simplificada. Esse atributo o posiciona como um candidato adequado para aplicações em dispositivos móveis.

A versão 5 da YOLO marca a primeira implementação nativa do PyTorch, destacando sua acessibilidade, treinamento eficiente e recursos de inferência, apresentando recursos oriundos de seu antecessor, a YOLOv4, como o SPP-NET. Ela também explora novos caminhos no aumento de dados por meio de metodologias como treinamento em mosaico e treinamento auto-antagônico (SAT). O YOLOv5 apresenta o PANet como uma alternativa à Feature Pyramid Network (FPN).

4.3. Métricas de Avaliação

A avaliação do desempenho da rede consiste na quantificação da eficácia de determinado modelo a fim de obter conclusões sobre a viabilidade de seu uso. A metodologia aplicada

para avaliar a rede YOLOv5 neste projeto gira em torno da matriz de confusão, na qual existem quatro valores principais: Verdadeiro Positivo (VP), Falso Positivo (FP), Falso Negativo (FN) e Verdadeiro Negativo (VN).

A partir destes valores é possível calcular as métricas que fornecem a compreensão do desempenho da rede. A precisão quantifica as previsões positivas calculando a proporção de verdadeiros positivos para a soma de verdadeiros positivos e falsos positivos. O *recall* mede a capacidade da rede de detectar positivos verdadeiros. É calculado como a proporção de verdadeiros positivos para a soma de verdadeiros positivos e falsos negativos. O *F1-score* leva em consideração falsos positivos e falsos negativos, consistindo na média harmônica entre precisão e *recall*.

5. Resultados e Discussão

Nesta seção são apresentados os resultados da construção do protótipo da aplicação, bem como do treinamento da rede utilizando a base de dados citada anteriormente.

5.1. Protótipo da Aplicação

A aplicação proposta pressupõe uma forma de controle de acesso (login) visto que as informações dos pacientes registrados são de caráter privado. O protótipo preparado inclui:

- Tela de login/cadastro;
- Tela inicial, com a lista de pacientes;
- Informações de cada paciente, abrangendo também seu histórico médico e sobre a evolução da úlcera;
- Tela de captura/visualização de imagem e tela de resultados, com as quais o modelo CNN irá interagir.

A Figura 1 apresenta as telas iniciais do aplicativo, sendo essas as primeiras com as quais o usuário interage. A Figura 2 apresenta detalhes mais aprofundados a respeito de cada paciente, que serão construídos a partir de dados externos e do registro das inferências da rede. A Figura 3 representa as telas relacionadas ao reconhecimento das lesões, sendo portanto o centro do projeto proposto.

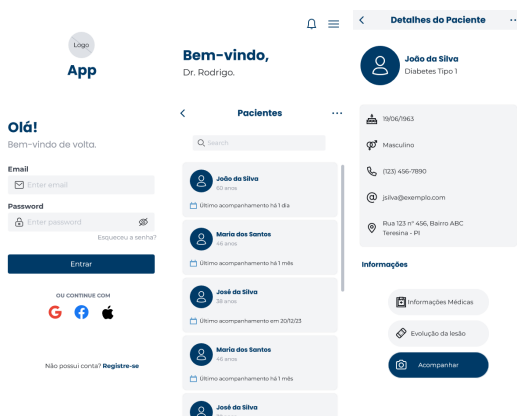


Figure 1. Telas de login, lista de pacientes e informações do paciente, da esquerda para a direita.

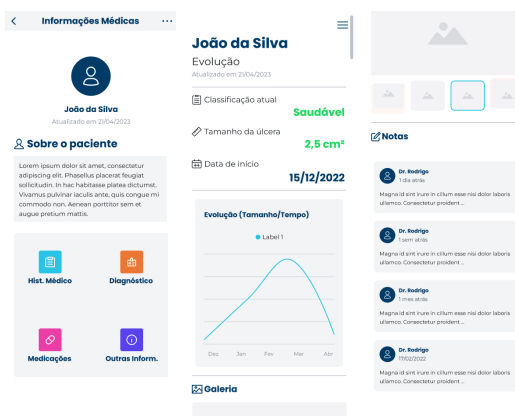


Figure 2. Tela com histórico médico do paciente e evolução (partes 1 e 2) da esquerda para a direita.

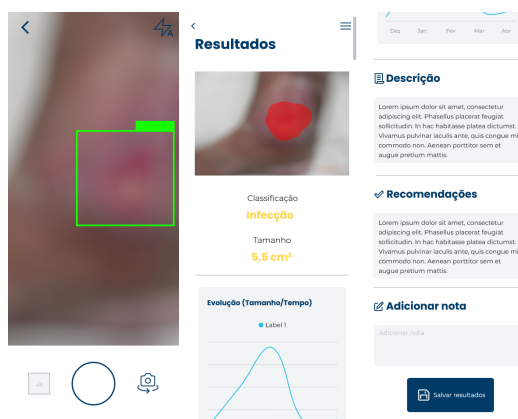


Figure 3. Tela de captura de imagem e resultados (partes 1 e 2) da esquerda para a direita.

5.2. Treinamento da Rede

Inicialmente, a rede YOLOv5 foi treinada utilizando inicialização randômica. A partir de então, os próximos treinos utilizavam os anteriores como checkpoint a fim de obter resultados melhores em menos tempo. Utilizando esta metodologia foram realizados 5 treinos consecutivos, modificando as imagens presentes nos conjuntos de treino, validação e teste a cada treino realizado.

Os resultados apresentados na Tabela 1 mostram que o modelo exibiu métricas positivas. Foi alcançado um *Mean Average Precision* de 86,20%, exemplificando a capacidade do modelo de localizar úlceras com precisão em imagens variadas. A precisão alcançou um valor de 84,82%, enquanto o *recall*, indicativo da sensibilidade do modelo, ficou em 83,82%. O *F1-score*, harmonizando precisão e *recall*, culminou em 84,31%, afirmando o desempenho equilibrado da rede.

CNN	A(%)	P(%)	R(%)	F(%)
YOLOv5	86.20 ± 1.69	84.82 ± 2.17	83.82 ± 0.57	84.31 ± 1.27

Table 1. Resultados obtidos no treinamento da rede representados pelas métricas escolhidas.

Na esquerda da Figura 4 é possível observar uma inferência bem-sucedida do modelo treinado. A imagem contém o pé de um paciente com uma lesão. O modelo detecta e localiza com precisão, cuja área é definida pela caixa delimitadora desenhada em torno dela.

A imagem da direita apresenta um cenário complexo que resultou em duas inferências. Uma delas identifica corretamente uma lesão, delineando com precisão seus limites. No entanto, a segunda inferência destaca um falso positivo, em que o modelo detecta erroneamente um objeto como uma lesão, resultando em uma caixa delimitadora que não corresponde a uma lesão real.



Figure 4. Na esquerda um exemplo de inferência bem sucedida da rede. Na direita, uma inferência correta e uma incorreta (Falso Positivo).

6. Conclusão e Trabalhos Futuros

O trabalho realizado representa um passo inicial significativo em direção à aplicação final desejada. Foram definidos os requisitos, padrões de projeto, arquitetura de software e tecnologias. A partir disso, foi construído um protótipo ilustrativo baseado em telas, facilitando a visualização do projeto.

Em seguida, foi selecionado e treinado um modelo de Rede Neural Convolutiva utilizando uma base abrangente de imagens. Os resultados do treinamento foram notavelmente positivos, com um Mean Average Precision de 86,20%, precisão de 84,82%, recall de 83,82% e um F1-Score de 84,31%.

É importante ressaltar que os objetivos esperados para este trabalho representam apenas uma fração inicial dos objetivos referentes à aplicação final que será desenvolvida futuramente. Em suma, o projeto estabelece as bases e fornece um caminho claro para trabalhos futuros e busca fazer avanços no campo do diagnóstico inteligente por imagem.

Para trabalhos futuros recomenda-se seguir com a construção da aplicação, implementando as camadas da arquitetura e integrando a rede treinada de forma que se faça o reconhecimento em tempo real em dispositivos móveis. Também é de grande interesse que se compare o desempenho da YOLOv5 com o de outros modelos a fim de garantir o uso de uma arquitetura tão eficiente quanto possível. Por fim, espera-se que seja possível treinar modelos que realizem a classificação e segmentação das úlceras, proporcionando assim um apoio ainda maior ao diagnóstico.

References

- Facebook (2023). React native getting started guide. <https://reactnative.dev/docs/getting-started>. Acesso em 23 mar. 2023.
- Galdran, A., Carneiro, G., and Ballester, M. A. G. (2021). Convolutional nets versus vision transformers for diabetic foot ulcer classification.
- Google (2023). Firebase documentation. <https://firebase.google.com/docs?hl=pt-br>. Acesso em 23 mar. 2023.
- Goyal, M., Reeves, N. D., Rajbhandari, S., Ahmad, N., Wang, C., and Yap, M. H. (2020). Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques. *Computers in Biology and Medicine*, 117:103616.
- Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., Michael, K., TaoXie, Fang, J., imyhxy, Lorna, Yifu), Wong, C., V, A., Montes, D., Wang, Z., Fati, C., Nadar, J., Laughing, UnglvKitDe, Sonck, V., tkianai, yxNONG, Skalski, P., Hogan, A., Nair, D., Strobel, M., and Jain, M. (2022). ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation.
- Ploderer, B., Brown, R., Seng, L., Lazzarini, P., and Van Netten, J. (2018). Promoting self-care of diabetic foot ulcers through a mobile application: A user-centred design and evaluation (preprint). *JMIR Diabetes*, 3.
- Pop, D.-P. and Samuel, A. A. (2014). Designing an mvc model for rapid web application development. *Procedia Engineering*, 69.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788.
- Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Unwin, N., Colagiuri, S., Guariguata, L., Motala, A. A., Ogurtsova, K., Shaw, J. E., Bright, D., and Williams, R. (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the international diabetes federation diabetes atlas. *Diabetes research and clinical practice*.
- Santos, E., Santos, F., Dallyson, J., Aires, K., Tavares, J., and Veras, R. (2022). Diabetic foot ulcers classification using a fine-tuned cnns ensemble. pages 282–287.
- Yap, M. H., Reeves, N., Boulton, A., Rajbhandari, S., Armstrong, D., Maiya, A. G., Najafi, B., Frank, E., and Wu, J. (2020). Diabetic foot ulcers grand challenge 2020.