

Modelo computacional preditivo para auxílio no diagnóstico de câncer de tireoide

Arthur Henrique Gomes Martins¹, Helen Cristina de Mattos Senefonte¹,
Iván Robert Enríquez Guzman²

¹Departamento de Computação – Universidade Estadual de Londrina (UEL)
Caixa Postal 10.011 – CEP 86057-970 – Londrina – PR – Brasil

²Departamento de Estatística - Universidade Federal do Espírito Santo (UFES)
Av. Fernando Ferrari, 514 - Goiabeiras – CEP 29075-910 Vitória – ES – Brasil

arthur.martins@uel.br, helen@uel.br, ivan.guzman@ufes.br

Abstract. *Thyroid cancer diagnoses are on the rise, and early detection is crucial for improving patient treatment outcomes. This necessitates the development of artificial intelligence models capable of accurately identifying malignant thyroid nodules and providing effective support to medical professionals. This study aims to implement a hybrid model that combines Vision Transformer (ViT) and Resnet, modifying activation functions to evaluate the impact of these changes on model performance. The research will conduct a comparative analysis of the results obtained from different model variations, as well as assess the accuracy of the original hybrid model, its modifications, and other state-of-the-art algorithms. The methodology includes the separate development of each algorithm, followed by their combined implementation and a detailed analysis of the results with the respective modifications. The study aims to identify which activation functions yield superior performance in the hybrid model, thereby enhancing the accuracy and efficiency of thyroid cancer detection through ultrasound imaging.*

Resumo. *O câncer de tireoide está em ascensão nos diagnósticos, e a sua detecção precoce é essencial para aumentar as chances de tratamento bem-sucedido dos pacientes. Isso demanda a criação de modelos de inteligência artificial que possam identificar com alta precisão nódulos malignos na tireoide e fornecer suporte eficaz aos médicos em suas decisões. O objetivo deste estudo é implementar um modelo híbrido que combine Vision Transformer (ViT) e Resnet, realizando modificações nas funções de ativação para avaliar como essas alterações influenciam o desempenho do modelo. A pesquisa promoverá uma análise comparativa entre os resultados obtidos pelos diferentes estilos do modelo, além de avaliar a acurácia do modelo híbrido original, suas modificações e outros algoritmos de ponta. Os métodos incluem o desenvolvimento separado de cada algoritmo, seguido pela implementação combinada e uma análise detalhada dos resultados com as devidas alterações. Espera-se que o estudo evidencie quais funções de ativação garantem um desempenho superior ao modelo híbrido, contribuindo para a melhoria da acurácia e eficiência na detecção de câncer de tireoide por meio de imagens de ultrassom.*

1. Introdução

O câncer na tireoide é um nódulo maligno que compromete a glândula do sistema endócrino. Atualmente, é o nono câncer mais frequente no mundo [5]. Sua incidência tem aumentado devido à maior detecção de certos tipos de câncer de tireoide, principalmente pelo uso de imagens na análise médica, como a ultrassonografia [5, 9].

A relevância do uso de imagens de ultrassom fica evidente em [11], em que pacientes diagnosticados com câncer de tireoide por meio desse exame apresentaram uma sobrevivência maior em comparação a diagnósticos feitos com outros tipos de imagens. Além disso, detectar o tumor precocemente é fundamental, pois as taxas de sobrevivência são significativamente maiores [9] e podem evitar a remoção total da glândula por meio da tireoidectomia.

Ao aliar questões médicas com o avanço da tecnologia, o *Computer-Aided Diagnosis* (CAD) torna-se imprescindível para uma maior acurácia no reconhecimento de patologias. Esse conceito refere-se ao uso de sistemas computacionais para contribuir na interpretação de imagens médicas, auxiliando na detecção, descrição e diagnóstico de anomalias [8]. Dessa forma, as análises médicas, que são subjetivas, contam com o benefício do computador para dar suporte as suas decisões.

Ademais, o aprendizado de máquina, juntamente com o aprendizado profundo, auxilia no desenvolvimento do CAD, pois analisa uma vasta quantidade de dados e identifica padrões que aprimoram a precisão do sistema. O uso desses algoritmos, especialmente os de aprendizado profundo, tem mostrado grande êxito na classificação e segmentação de imagens médicas [10]. Esta sinergia entre CAD e aprendizado de máquina garante não só um aprimoramento da precisão diagnóstica, mas também acelera o processo, permitindo diagnósticos mais rápidos e intervenções precoces, o que é crucial, como observado no câncer de tireoide.

Dado o impacto significativo do CAD na precisão e eficiência dos diagnósticos médicos, é necessário buscar continuamente o aprimoramento desses sistemas, especialmente na glândula tireoide. Para isso, deve-se investir no refinamento de algoritmos de aprendizado de máquina. Esse é o caso do modelo ViT híbrido [13], no qual os autores combinaram duas técnicas robustas e consolidadas de aprendizado profundo em reconhecimento de imagem para gerar o diagnóstico de câncer de tireoide por meio de imagens de ultrassom.

Dessa forma, o objetivo deste trabalho de conclusão de curso é implementar um modelo híbrido de *Vision Transformer* (ViT) [7] com *Resnet* [12], com ênfase na modificação das funções de ativação para avaliar seu impacto no desempenho do modelo. O estudo visa comparar os resultados obtidos antes e após as mudanças, buscando identificar quais funções de ativação melhoram a precisão na detecção de câncer de tireoide em imagens médicas.

A Seção 2 apresenta a revisão do estado da arte e conceitos importantes para o trabalho. Nas Seções 3 e 4 estão os objetivos a serem alcançados e os procedimentos metodológicos para atingi-los. A Seção 5 exhibe o cronograma e, por fim, a Seção 6 apresenta as contribuições e os resultados esperados deste trabalho de conclusão de curso.

2. Fundamentação Teórico-Metodológica e Estado da Arte

Como mencionado anteriormente, a detecção precoce do câncer na tireoide é crucial para aumentar as chances de tratamento bem-sucedido e sobrevida dos pacientes. Isso torna necessário o desenvolvimento de modelos de inteligência artificial capazes de identificar com a maior precisão possível um nódulo maligno na tireoide. Tradicionalmente, os pesquisadores trabalham com imagens de ultrassom como fonte de dados para suas pesquisas, mas há também estudos que utilizam outras abordagens.

2.1. Estado da Arte

Pathack *et al.* em [19] e Do *et al.* em [6] são pesquisas que empregaram como fonte de dados imagens diferentes da ultrassonografia, ambos empregando imagens microscópicas. O primeiro associou uma rede neural com um sensor fotônico, enquanto o segundo desenvolveu um modelo baseado na rede neural convolucional (*CNN*) *Inception v3*, classificando o tipo de tireoide em seis tumores distintos. A partir do próximo parágrafo, são citados artigos que utilizaram imagens de ultrassom para as suas pesquisas.

Possuir uma base de dados com quantidades significativas é um fator determinante em inteligência artificial, embora no estudo de detecção de câncer na tireoide os dados sejam escassos. Contudo, essa limitação não impede que as pesquisas obtenham bons resultados. Yang *et al.* em [27] evidencia esse fato, propondo um método com melhorias baseado no modelo de detecção de imagem YOLO (*You Only Look Once*) com um conjunto de dados de aproximadamente 200 imagens, alcançando uma acurácia de 95,3%. Outros estudos, como [25], [14] e [18] obtiveram acurácias acima de 90% com um conjunto de dados de tamanho semelhante, sendo o primeiro, que utiliza filtros quânticos associados a uma rede neural convolucional, o maior com 97,72% na identificação de tumores malignos e benignos.

Em contrapartida, [28] propôs o desenvolvimento da rede neural convolucional VGG16 com algumas alterações, obtendo um sucesso inferior aos outros mencionados, mas ainda superior ao modelo original. Já as pesquisas [1] e [17] conseguiram utilizar grandes quantidades de dados. Em [1], cerca de três mil imagens foram usadas para o treinamento do *Mask R-CNN* desenvolvido pelos autores, enquanto o segundo utilizou significativas 15 mil imagens. Neste último, os autores combinaram duas *CNNs* para extração de características e classificação binária. Todavia, mesmo com um grande conjunto de imagens, as duas pesquisas não alcançaram acurácias relevantes para a detecção do câncer.

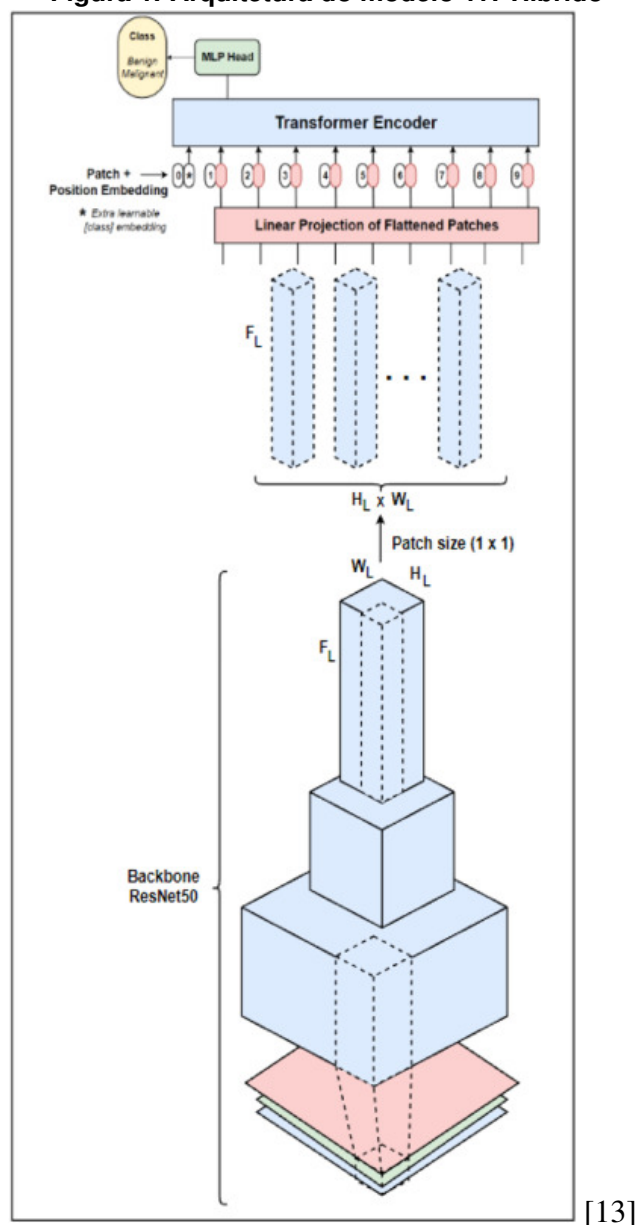
2.2. Deep Learning

Deep Learning, ou aprendizado profundo, é uma especialização dentro do aprendizado de máquina que utiliza redes neurais profundas para modelar padrões complexos em grandes volumes de dados. Este campo tem se destacado em áreas como processamento de imagens e compreensão de linguagem natural onde dados brutos são processados por várias camadas para extrair características de nível superior [15]. Essa técnica permite a melhora da capacidade dos modelos de realizar previsões precisas e reconhecer padrões complexos em grandes conjuntos dados.

Assim, a fim de solucionar a classificação de imagens de ultrassom de tireoide, os pesquisadores Feres Jerbi, Noura Aboudi e Nawres Khelifa em [13] desenvolveram um

modelo de aprendizado profundo que combina *Vision Transformer* (ViT) com a rede neural convolucional *Resnet50*, chamado de ViT Híbrido. Sua arquitetura, como é possível ver na Figura 1, funciona de forma que a *Resnet* atue como o esqueleto, extraíndo as características mais importantes da imagem de ultrassom. Essas características transformadas em um vetor serão a entrada do ViT, que fará o reconhecimento de padrões da sua arquitetura e então classificará o tumor como maligno ou benigno. Diante disso, é essencial o conhecimento sobre redes neurais convolucionais e *Vision Transformers*.

Figura 1. Arquitetura do modelo ViT Híbrido



[13]

2.2.1. Redes Neurais Convolucionais

Uma rede neural convolucional (CNN) é definida em [21] como um tipo de rede neural profunda projetada para processamento de dados visuais, como imagens. As CNNs utili-

zam a operação de convolução para detectar padrões locais, aplicando filtros em pequenas regiões da imagem. A arquitetura típica inclui camadas de convolução e *pooling*, repetidas várias vezes, seguidas por camadas totalmente conectadas que realizam a classificação final dos dados visuais.

Pesquisas já citadas [1, 6, 25, 28, 17] utilizaram como base para seus modelos a CNN em essência, mas com alterações em algumas partes de suas arquiteturas. Além disso, [22], com a arquitetura baseada no algoritmo *EfficientNet*, e [16], com o *Faster R-CNN*, também usaram esse tipo de rede neural profunda para a classificação de nódulos de tireoide, resultando em uma ótima acurácia, sobretudo o primeiro.

2.2.2. ResNet

A *Resnet* [12], ou Rede Residual, é uma arquitetura de redes neurais convolucionais projetada para treinar redes muito profundas de forma eficaz, introduzindo o conceito de aprendizado residual. Em vez de aprender diretamente as funções de mapeamento, a ResNet aprende funções residuais utilizando conexões de atalho que pulam uma ou mais camadas. Isso permite o treinamento de redes com muitas camadas, sem perda de acurácia. Existem várias variantes desse modelo, e a utilizada no ViT Híbrido possui 50 camadas de convolução.

Pavithra *et al.* [20] e Alghanimi *et al.* [2] utilizaram redes residuais para a classificação de câncer na tireoide, sendo o segundo responsável por definir o tipo de nódulo cancerígeno. Além deles, Amgad [3] comparou os resultados de arquiteturas de CNN, incluindo a *VGG16* [28], e verificou que a *ResNet50* apresentou um melhor desempenho, com acurácia de 92%.

2.2.3. Vision Transformers

Introduzido por Alexey Dosovitskiy em [7], o *Vision Transformer* (ViT) foi desenvolvido a partir do *Transformer* de processamento de linguagem natural [26] e surgiu como uma alternativa às arquiteturas convolucionais para o reconhecimento de imagens. Nesse modelo, a imagem é dividida em fragmentos chamados *patches*, que são alimentados a um codificador que busca relações entre os fragmentos. Por fim, a arquitetura termina com uma rede neural para a classificação da imagem.

Esse robusto modelo teve um grande impacto no processamento de imagens com aprendizado profundo. Em [23], os autores enaltecem as qualidades do *Vision Transformer* em comparação a modelos consolidados como as CNNs. Além disso, o artigo também destaca a combinação de outros modelos com o ViT e sua capacidade de classificação de imagens médicas. Isso é evidente também em [4, 24], onde os autores utilizaram modelos baseados em ViT para a detecção do câncer de tireoide, obtendo boa acurácia na classificação.

3. Objetivos

Esse trabalho tem como objetivo implementar, de forma eficiente e precisa, um modelo que poderá ser utilizado no auxílio de detecção de tumores cancerígenos na glândula

tireoide através de:

1. Implementar o modelo híbrido proposto por [13];
2. Investigar e adaptar as funções de ativação do modelo, a fim de avaliar o impacto dessas alterações na performance do modelo;
3. Promover um estudo comparativo entre os resultados obtidos pelos diferentes estilos do modelo;
4. Avaliar a precisão do diagnóstico de câncer de tireoide do modelo híbrido original, suas alterações e de outros algoritmos do estado da arte.

4. Procedimentos metodológicos/Métodos e técnicas

Primeiramente, será realizada uma revisão bibliográfica sobre os algoritmos utilizados para a detecção do câncer de tireoide, visando identificar quais foram os métodos mais usados e eficazes, além do estudo da *Resnet* [12] e do *Vision Transformer* [7], modelos empregados no modelo híbrido, a fim de compreender suas arquiteturas e a maneira como podem ser combinadas.

Após isso, cada algoritmo será desenvolvido separadamente para a problemática, com o intuito de observar seus comportamentos. Em seguida, será feita a implementação dos dois modelos de forma combinada e, logo após, as funções de ativação serão alteradas para realizar um estudo comparativo dos resultados obtidos e gerar análises sobre as mudanças.

5. Cronograma de Execução

Atividades:

1. Revisão bibliográfica sobre os métodos utilizados para detectar o câncer de tireoide;
2. Desenvolvimento dos algoritmos separadamente para a problemática;
3. Implementação e validação do modelo híbrido;
4. Investigação e adaptação das funções de ativação do modelo;
5. Análise dos resultados obtidos;
6. Escrita de artigos para submissão em conferências;
7. Escrita da versão preliminar do TCC;
8. Escrita da versão final do TCC;

Tabela 1. Cronograma de Execução

	ago	set	out	nov	dez	jan	fev
Atividade 1	•	•	•	•			
Atividade 2	•	•					
Atividade 3	•	•	•	•			
Atividade 4			•	•			
Atividade 5		•	•	•	•	•	
Atividade 6			•	•	•	•	•
Atividade 7	•	•	•				
Atividade 8			•	•	•	•	•

6. Contribuições e/ou Resultados esperados

Ao final do trabalho, são esperado os seguintes pontos:

- Desenvolvimento do ViT híbrido capaz de detectar câncer de tireoide com eficiência;
- Identificação de quais funções de ativação proporcionam melhores resultados no contexto de câncer de tireoide;
- Comparação entre a precisão diagnóstica do modelo híbrido aprimorado e a versão original, demonstrando as melhorias alcançadas;
- Análise comparativa do modelo desenvolvido com algoritmos propostos pelo estado da arte;
- Desenvolvimento maior dos sistemas de *Computer-aided Diagnosis* e incentivo às pesquisas na busca pelo conhecimento nessa área.

7. Espaço para assinaturas

Londrina, 29/07/2024.



Aluno



Orientador

Referências

- [1] Fatemeh Abdolali, Jeevesh Kapur, Jacob L. Jaremko, Michelle Noga, Abhilash R. Harendranathan, and Kumaradevan Punithakumar. Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks. *Computers in Biology and Medicine*, 122:103871, 2020.
- [2] Ghufuran B. Alghanimi, Hadeel K. Aljobouri, and Khaleel Akeash Al-shimmari. Cnn and resnet50 model design for improved ultrasound thyroid nodules detection. In *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETISIS)*, pages 1000–1004, 2024.
- [3] Nadeen Amgad, Hadiy Haitham, Mona Alabarak, and Ammar Mohammed. Enhancing thyroid cancer diagnosis through a resilient deep learning ensemble approach. In *2024 6th International Conference on Computing and Informatics (ICCI)*, pages 195–202, 2024.
- [4] Namu Baima, Tianfu Wang, Chong-Ke Zhao, Siping Chen, Chen Zhao, and Baiying Lei. Dense swin transformer for classification of thyroid nodules. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pages 1–4, 2023.
- [5] Laura Boucai, Mark Zafereo, and Maria E. Cabanillas. Thyroid Cancer: A Review. *JAMA*, 331(5):425–435, 02 2024.
- [6] Thanh-Ha Do and Hoa Nguyen Khanh. Supporting thyroid cancer diagnosis based on cell classification over microscopic images. In *2022 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, pages 1–5, 2022.

- [7] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale, 2021.
- [8] B. Erickson and B. Bartholmai. Computer-aided detection and diagnosis at the start of the third millennium. *Journal of Digital Imaging*, 2002. Cancer-related cognitive impairment.
- [9] S. Filetti, C. Durante, D. Hartl, S. Leboulleux, L.D. Locati, K. Newbold, M.G. Papotti, and A. Berruti. Thyroid cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up††approved by the esmo guidelines committee: February 2008, last update september 2019. this publication supersedes the previously published version—ann oncol 2012; 23(suppl 7): vii110–vii119. *Annals of Oncology*, 30(12):1856–1883, 2019. Cancer-related cognitive impairment.
- [10] Arunim Garg and Vijay Mago. Role of machine learning in medical research: A survey. *Computer Science Review*, 40:100370, 2021.
- [11] Megan R. Haymart, David Reyes-Gastelum, Elaine Caoili, Edward C. Norton, and Mousumi Banerjee. The Relationship Between Imaging and Thyroid Cancer Diagnosis and Survival. *The Oncologist*, 25(9):765–771, 05 2020.
- [12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.
- [13] Feres JERBI, Noura ABOUDI, and Nawres KHLIFA. Automatic classification of ultrasound thyroids images using vision transformers and generative adversarial networks. *Scientific African*, 20:e01679, 2023.
- [14] Xhitij A. Kesarkar and K.V. Kulhalli. Thyroid nodule detection using artificial neural network. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, pages 11–15, 2021.
- [15] Yann LeCun, Y. Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436–44, 05 2015.
- [16] Hailiang Li, Jian Weng, Yujian Shi, Wanrong Gu, Yijun Mao, Yonghua Wang, Weiwei Liu, and Jiajie Zhang. An improved deep learning approach for detection of thyroid papillary cancer in ultrasound images. *Scientific Reports*, 8, 04 2018.
- [17] Jinlian Ma, Fa Wu, Jiang Zhu, Dong Xu, and Dexing Kong. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. *Ultrasonics*, 73:221–230, 2017.
- [18] Xuesi Ma and Lina Zhang. Diagnosis of thyroid nodules based on image enhancement and deep neural networks. *Computational Intelligence and Neuroscience*, 2022(1):5582029, 2022.
- [19] Ranjeet Kumar Pathak, Sumita Mishra, and Preeta Sharan. Thyroid cancer detection using artificial neural network and photonic sensor. In *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, volume 10, pages 1298–1303, 2023.

- [20] S Pavithra, G Yamuna, and R Arunkumar. Deep learning method for classifying thyroid nodules using ultrasound images. In *2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*, pages 1–6, 2022.
- [21] Zakaria Rguibi, Abdelmajid Hajami, Dya Zitouni, Amine Elqaraoui, and Anas Bedraoui. Cxai: Explaining convolutional neural networks for medical imaging diagnostic. *Electronics*, 11(11), 2022.
- [22] Md Faizan Shakeel, Mudassir Hasan Khan, and Yusuf Uzzaman Khan. Deep learning empowering diagnosis of thyroid nodule malignancy through ultrasound imaging. In *2023 International Conference on Recent Advances in Science and Engineering Technology (ICRASET)*, pages 1–6, 2023.
- [23] Fahad Shamshad, Salman Khan, Syed Waqas Zamir, Muhammad Haris Khan, Munawar Hayat, Fahad Shahbaz Khan, and Huazhu Fu. Transformers in medical imaging: A survey. *Medical Image Analysis*, 88:102802, 2023.
- [24] Jiawei Sun, Bobo Wu, Tong Zhao, Liugang Gao, Kai Xie, Tao Lin, Jianfeng Sui, Xiaoqin Li, Xiaojin Wu, and Xinye Ni. Classification for thyroid nodule using vit with contrastive learning in ultrasound images. *Computers in Biology and Medicine*, 152:106444, 2023.
- [25] Gunti Swathi, Ali Altalbe, and R. Prasanna Kumar. Qucnet: Quantum-inspired convolutional neural networks for optimized thyroid nodule classification. *IEEE Access*, 12:27829–27842, 2024.
- [26] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2023.
- [27] Daqing Yang, Jianfu Xia, Rizeng Li, Wencai Li, Jisheng Liu, Rongjian Wang, Dong Qu, and Jie You. Automatic thyroid nodule detection in ultrasound imaging with improved yolov5 neural network. *IEEE Access*, 12:22662–22670, 2024.
- [28] Yi-Cheng Zhu, Peng-Fei Jin, Jie Bao, Quan Jiang, and Ximing Wang. Thyroid ultrasound image classification using a convolutional neural network. *Annals of Translational Medicine*, 9, 10 2021.