UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
CURSO DE GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO

Stephan Krug

Triagem automatizada de alterações macroscópicas placentárias com redes neurais convolucionais para auxílio na solicitação de exames anatomopatológicos

Florianópolis
2025

Stephan Krug

Triagem automatizada de alterações macroscópicas placentárias com redes neurais convolucionais para auxílio na solicitação de exames anatomopatológicos

Trabalho de Conclusão de Curso do Curso de Graduação em Ciências da Computação do Centro Tecnológico da Universidade Federal de Santa Catarina para a obtenção do título de bacharel em ciências da computação.
Orientador: Prof. Aldo von Wangenheim, Dr. rer. nat.

Florianópolis
2025

Stephan Krug

Triagem automatizada de alterações macroscópicas placentárias com redes neurais convolucionais para auxílio na solicitação de exames anatomopatológicos

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de "bacharel em ciências da computação" e aprovado em sua forma final pelo Curso de Graduação em Ciências da Computação.

Florianópolis, 7 de julho de 2025.

Profa. Lúcia Helena Martins Pacheco, Dra.
Coordenadora do Curso

Banca Examinadora:

Prof. Aldo von Wangenheim, Dr. rer. nat.
Orientador

Bibiana Quatrin Tiellet da Silva, Ma.
Avaliadora
Universidade Federal do Rio Grande do Sul

Beatriz Silva Lopes, B.
Avaliadora
Instituto Federal Catarinense

AGRADECIMENTOS

Agradeço sobretudo à minha família: minha mãe, Miriam, meu pai, Gerson, minha irmã, Yasmin, e meus avós, Edeltraut, Fridel, Giesela e Albert (*in memoriam*), por todo o amor e por me proporcionarem tudo que fosse preciso para que eu pudesse trilhar este caminho. Sem vocês, nada disso seria possível.

À minha namorada, Ana, sou grato pelo amor, apoio e por estar ao meu lado em todos os momentos, tanto nos felizes quanto nos desafiadores da graduação e da vida. Sua presença me inspirou a ser melhor e a seguir sempre em frente.

Aos meus amigos e colegas, agradeço por compartilharem essa jornada comigo e por contribuírem com aprendizados e momentos que tornaram essa viagem mais leve.

Expresso também minha gratidão ao Aldo, pela orientação e recomendações durante o desenvolvimento deste trabalho. Agradeço igualmente aos participantes do grupo de pesquisa e membros da banca, Beatriz e Bibiana, por me acolherem e auxiliarem na condução do estudo. Seus ensinamentos levarei comigo pela vida inteira.

Aos meus colegas de trabalho nas empresas SC Clouds e V360, sou grato pelo apoio, pelas experiências e por me tornarem um profissional melhor.

Por fim, aos professores do curso de Ciências da Computação da UFSC pelas disciplinas desafiadoras e de alta qualidade, que me ensinaram e fizeram apaixonar ainda mais por esta ciência artística que move o mundo.

RESUMO

Patologias placentárias estão associadas a diversos desfechos perinatais desfavoráveis, sendo responsáveis por até 65% dos casos de óbito fetal. O diagnóstico preciso é realizado por meio de exame anatomopatológico; no entanto, como grande parte dos partos ocorre sem intercorrências e, aliado a questões de custo e logística, sua aplicação a todos os partos torna-se desnecessária e inviável. Por essa razão, a seleção dos casos que demandam o exame depende de uma análise macroscópica criteriosa e fundamentada, que, embora eficiente, pode ser trabalhosa e suscetível a subjetividades. Este trabalho propõe o uso de redes neurais convolucionais baseadas em aprendizado por transferência (*transfer learning*) como suporte automatizado à triagem de padrões morfológicos em placentas humanas. Foram empregados os modelos pré-treinados ResNet-18, ResNet-34, EfficientNet-B0 e EfficientNet-B4 para classificar tanto a presença ou ausência de alterações morfológicas quanto o tipo de anomalia presente, utilizando imagens de placentas provenientes do serviço de Obstetrícia do HU-UFSC/EBSERH e de outras fontes. Para contornar limitações de tamanho e desbalanceamento do conjunto de dados, aplicaram-se técnicas de *data augmentation*. Os modelos foram avaliados pelas métricas de acurácia, precisão, *recall* e F1-Score, e os resultados demonstraram elevado desempenho: na primeira tarefa, os modelos ResNet-18 com *data augmentation* e EfficientNet-B4 com e sem *data augmentation* atingiram 100% em todas as métricas, enquanto na segunda tarefa o modelo EfficientNet-B0 obteve acurácia de 97,4% e F1-Score de 91,5%. Esses resultados confirmam a viabilidade do uso de aprendizado profundo para classificação de anomalias placentárias, evidenciando o potencial de modelos pré-treinados como ferramentas de apoio à decisão clínica.

Palavras-chave: Placenta. Redes Neurais Convolucionais. Aprendizado por transferência. Classificação de imagens. Análise macroscópica de placentas. Alterações preditivas de doenças em placentas.

ABSTRACT

Placental pathologies are associated with various adverse perinatal outcomes and are responsible for up to 65% of fetal deaths. Accurate diagnosis is performed through anatomopathological examination; however, since most births occur without complications and, combined with cost and logistical issues, its application to all births becomes unnecessary and unfeasible. For this reason, the selection of cases requiring the examination depends on a careful and well-founded macroscopic analysis, which, although efficient, can be labor-intensive and susceptible to subjectivity. This study proposes the use of convolutional neural networks based on transfer learning as an automated support for screening morphological patterns in human placentas. Pre-trained models ResNet-18, ResNet-34, EfficientNet-B0, and EfficientNet-B4 were employed to classify both the presence or absence of morphological alterations and the type of anomaly present, using images of placentas from the Obstetrics Department of the University Hospital Polydoro Ernani de São Thiago (HU-UFSC/EBSERH) and other sources. To address dataset size and imbalance limitations, data augmentation techniques were applied. The models were evaluated using accuracy, precision, recall, and F1-Score metrics, and the results demonstrated high performance: in the first task, ResNet-18 with data augmentation and EfficientNet-B4 with and without data augmentation achieved 100% across all metrics, while in the second task, EfficientNet-B0 achieved an accuracy of 97.4% and an F1-Score of 91.5%. These results confirm the feasibility of using deep learning for placental anomaly classification, highlighting the potential of pre-trained models as decision-support tools in clinical practice.

Keywords: Placenta. Convolutional Neural Networks. Transfer learning. Image classification. Macroscopic placental analysis. Disease-predictive placental alterations.

ZUSAMMENFASSUNG

Plazentapathologien sind mit verschiedenen ungünstigen perinatalen Ergebnissen assoziiert und verantwortlich für bis zu 65% der Fälle von fetalem Tod. Die präzise Diagnose erfolgt durch eine anatomopathologische Untersuchung; jedoch wird ihre Anwendung auf alle Geburten aufgrund der Tatsache, dass ein großer Teil der Entbindungen ohne Komplikationen verläuft und in Verbindung mit Kosten- und logistischen Fragen unnötig und nicht durchführbar. Aus diesem Grund hängt die Auswahl der Fälle, die eine Untersuchung erfordern, von einer sorgfältigen und fundierten makroskopischen Analyse ab, die zwar effizient ist, jedoch aufwendig sein kann und anfällig für Subjektivität ist. Diese Arbeit schlägt die Verwendung von konvolutionalen neuronalen Netzen auf der Grundlage von Transferlernen (*transfer learning*) als automatisierte Unterstützung für das Screening morphologischer Änderungen in menschlichen Plazenten vor. Es wurden die vortrainierten Modelle ResNet-18, ResNet-34, EfficientNet-B0 und EfficientNet-B4 eingesetzt, um sowohl das Vorhandensein oder Fehlen morphologischer Veränderungen als auch deren Art zu klassifizieren. Die Bilder stammten aus dem Geburtsdienst des Universitätsklinikums HU-UFSC/EBSERH und aus anderen Quellen. Um Einschränkungen hinsichtlich der Datensatzgröße und des Ungleichgewichts zu überwinden, wurden Data-Augmentation-Techniken angewendet. Die Modelle wurden anhand der Metriken Genauigkeit (accuracy), Präzision (precision), Sensitivität (recall) und F1-Score bewertet. Die Ergebnisse zeigten eine hohe Leistungsfähigkeit: In der ersten Aufgabe erreichten ResNet-18 mit Data Augmentation sowie EfficientNet-B4 mit und ohne Data Augmentation in allen Metriken 100%, während bei der zweiten Aufgabe das Modell EfficientNet-B0 eine Genauigkeit von 97.4% und einen F1-Score von 91.5% erzielte. Diese Ergebnisse bestätigen die Machbarkeit des Einsatzes von Deep Learning zur Klassifizierung plazentarer Anomalien und verdeutlichen das Potenzial vortrainierter Modelle als Werkzeuge zur Unterstützung klinischer Entscheidungsprozesse.

Schlüsselwörter: Plazenta. Konvolutionale neuronale Netze. Transferlernen. Bildklassifikation. Makroskopische Plazentaanalyse. Prädiktive Veränderungen von Plazentaerkrankungen.

RESUMEN

Las patologías placentarias están asociadas a diversos desenlaces perinatales adversos y son responsables de hasta el 65% de las muertes fetales. El diagnóstico preciso se realiza mediante examen anatomopatológico; sin embargo, dado que gran parte de los partos ocurre sin complicaciones y, sumado a cuestiones de costo y logística, su aplicación a todos los partos se vuelve innecesaria e inviable. Por esta razón, la selección de los casos que requieren el examen depende de un análisis macroscópico cuidadoso y fundamentado, que, aunque eficiente, puede ser laborioso y susceptible a subjetividades. Este estudio propone el uso de redes neuronales convolucionales basadas en aprendizaje por transferencia (*transfer learning*) como apoyo automatizado para la clasificación de patrones morfológicos en placentas humanas. Se emplearon los modelos preentrenados ResNet-18, ResNet-34, EfficientNet-B0 y EfficientNet-B4 para clasificar tanto la presencia o ausencia de alteraciones morfológicas como el tipo de anomalía presente, utilizando imágenes de placentas provenientes del servicio de Obstetricia del Hospital Universitario Polydoro Ernani de São Thiago (HU-UFSC/EBSERH) y de otras fuentes. Para superar las limitaciones de tamaño y desbalance del conjunto de datos, se aplicaron técnicas de *data augmentation*. Los modelos fueron evaluados mediante las métricas de exactitud, precisión, *recall* y F1-Score, y los resultados demostraron un alto rendimiento: en la primera tarea, los modelos ResNet-18 con *data augmentation* y EfficientNet-B4 con y sin *data augmentation* alcanzaron el 100% en todas las métricas, mientras que en la segunda tarea el modelo EfficientNet-B0 obtuvo una exactitud del 97,4% y un F1-Score del 91,5%. Estos resultados confirman la viabilidad del uso del aprendizaje profundo para la clasificación de anomalías placentarias, destacando el potencial de los modelos preentrenados como herramientas de apoyo a la decisión clínica.

Palabras clave: Placenta. Redes neuronales convolucionales. Aprendizaje por transferencia. Clasificación de imágenes. Análisis macroscópico placentario. Alteraciones placentarias predictivas de enfermedades.

LISTA DE FIGURAS

Figura 1 – Um esquema representativo de uma rede neural artificial (DAS; SADIQ; MIRJALILI, 2025). 19
Figura 2 – Um esquema representativo de uma rede neural convolucional (LECUN; BOTTOU *et al.*, 1998). 20
Figura 3 – Um esquema representativo de uma camada de convolução (IBM, 2025). 21
Figura 4 – Um esquema do funcionamento de uma *pooling layer* (YINGGE; ALI; LEE, 2020). 22
Figura 5 – Um esquema representativo de uma *fully connected layer* (HOPE; RESHEFF; LIEDER, 2018). 23
Figura 6 – Um comparativo das funções de ativação sigmoid, tanh, ReLU e Leaky ReLU (KUO; LIN; HU, 2020). 24
Figura 7 – Exemplos de aumentação de dados (MUREL; KAVLAKOGLU, 2024). 26
Figura 8 – Ilustração da variação da taxa de aprendizado e do momentum com a *one cycle policy* (ROMERO *et al.*, 2019). 27
Figura 9 – Representação gráfica do modelo ResNet (HE *et al.*, 2016). 28
Figura 10 – Representação gráfica dos conceitos de profundidade, largura e resolução para o escalonamento composto (TAN; LE, 2019). 29
Figura 11 – Imagens de uma placenta expondo a porção materna (esquerda) e fetal (direita). 32
Figura 12 – Exemplo de uma placenta morfologicamente alterada, com um nó verdadeiro no cordão umbilical. 35
Figura 13 – Resumo da metodologia de revisão sistemática da literatura, descrito por Kitchenham (PEREIRA DA SILVA, 2009). 37
Figura 14 – Arquitetura do PlacentaNet (CHEN; WU *et al.*, 2019). 40
Figura 15 – Arquitetura expandida (CHEN; ZHANG *et al.*, 2020). 42
Figura 16 – *Pipeline* MSL proposto (ZHANG *et al.*, 2020). 43
Figura 17 – Diagrama representativo do método VLC proposto (PAN *et al.*, 2022). 43
Figura 18 – Diagrama ilustrando a diferença entre o método proposto (esquerda) e o método VLC tradicional (direita) (PAN *et al.*, 2023). 44
Figura 19 – Imagens de exemplo de cada classe de placentas. 47

Figura 20 – Matriz de confusão para as configurações que atingiram um F1-Sore de **1.000**. Arquiteturas/configurações: ResNet-18 (NoDA, BS=8), EfficientNet-B4 (NoDA, BS=8), ResNet-18 (DA, BS=8), EfficientNet-B4 (DA, BS=8), ResNet-34 (NoDA, BS=16), EfficientNet-B4 (NoDA, BS=16), ResNet-18 (DA, BS=16), ResNet-34 (DA, BS=16), EfficientNet-B0 (DA, BS=16), EfficientNet-B4 (DA, BS=16), ResNet-18 (NoDA, BS=32), EfficientNet-B4 (DA, BS=32), ResNet-18 (DA, BS=32), ResNet-34 (DA, BS=32), EfficientNet-B4 (DA, BS=32). . . . 53
Figura 21 – Matriz de confusão para as configurações que atingiram um F1-Sore de **0.884**. Arquiteturas/configurações: ResNet-34 (NoDA, BS=8), ResNet-34 (DA, BS=8), EfficientNet-B0 (DA, BS=8), ResNet-18 (NoDA, BS=16), EfficientNet-B0 (NoDA, BS=16), EfficientNet-B0 (NoDA, BS=32), ResNet-34 (DA, BS=32). 56
Figura 22 – Matriz de confusão para as configurações que atingiram um F1-Score de **0.839**. Arquiteturas/configurações: EfficientNet-B0 (NoDA, BS=32), ResNet-18 (NoDA, BS=32). 56
Figura 23 – Matriz de confusão para a arquitetura ResNet-18 na classificação do tipo de alteração em placentas. 58
Figura 24 – Matriz de confusão para a arquitetura ResNet-34 na classificação do tipo de alteração em placentas. 59
Figura 25 – Matriz de confusão para a arquitetura EfficientNet-B0 na classificação do tipo de alteração em placentas. 60
Figura 26 – Matriz de confusão para a arquitetura EfficientNet-B4 na classificação do tipo de alteração em placentas. 61
Figura 27 – Exemplos de visualizações Grad-CAM aplicadas a diferentes placentas no conjunto de teste, destacando as regiões de maior importância para a classificação do tipo de alteração morfológica. 62

LISTA DE TABELAS

Tabela 1 – Divisão das imagens do dataset por tipo de alteração. 46
Tabela 2 – Distribuição das imagens de placentas alteradas e não alteradas entre os conjuntos de treino, validação e teste 48
Tabela 3 – Distribuição das imagens por tipo de alteração entre os conjuntos de treino, validação e teste. 51
Tabela 4 – Resultados experimentais na fase de validação, com e sem data augmentation. 54
Tabela 5 – Resultados experimentais na fase de teste, com e sem data augmentation. 55
Tabela 6 – Resultados experimentais na fase de validação. 57
Tabela 7 – Resultados experimentais na fase de teste. 58

LISTA DE ABREVIATURAS E SIGLAS

HU-UFSC/EBSERH	Hospital Universitário Polydoro Ernani de São Thiago da Universidade Federal de Santa Catarina
CNN	*Convolutional Neural Network* (Rede Neural Convolucional)
IA	Inteligência Artificial
NLP	*Natural Language Processing* (Processamento de Linguagem Natural)
ML	*Machine Learning* (Aprendizado de Máquina)
DL	*Deep Learning* (Aprendizado Profundo)
Tanh	Tangente Hiperbólica
ReLU	*Rectified Linear Unit* (Unidade Linear Retificada)
MSE	*Mean Squared Error* (Erro Quadrático Médio)
SGD	*Stochastic Gradient Descent* (Descida de Gradiente Estocástico)
RMSProp	*Root Mean Square Propagation* (Propagação de Valor Quadrático Médio)
Adam	*Adaptive Moment Estimation* (Estimativa de Momento Adaptativo)
MBConv	*Mobile Inverted Bottleneck Convolution* (Convolução de Gargalo Móvel Invertido)
SE	*Squeeze-and-Excitation* (Compressão e Excitação)
MSL	*Multi-Region Saliency-Aware* (Reconhecimento de Saliência Multirregional)
VLC	*Vision Language Contrastive Learning* (Aprendizado Contrastivo de Visão e Linguagem)
LR	*Learning Rate* (Taxa de Aprendizado)
GPU	*Graphics Processing Unit* (Unidade de Processamento Gráfico)
HSV	*Hue-Saturation-Value* (Matiz, Saturação e Valor)
ViT	*Vision Transformers* (Transformador de Visão)

SUMÁRIO

1 INTRODUÇÃO 15
1.1 JUSTIFICATIVA 16
1.2 OBJETIVOS 16
1.2.1 Objetivo Geral 16
1.2.2 Objetivos Específicos 16
2 CONCEITOS RELACIONADOS 17
2.1 VISÃO COMPUTACIONAL 17
2.1.1 Classificação de imagens 17
2.2 INTELIGÊNCIA ARTIFICIAL (IA) 18
2.2.1 Aprendizado de máquina (*machine learning* - ML) 18
2.2.2 Aprendizado profundo (*deep learning* - DL) 18
2.2.3 Redes neurais artificiais (*artificial neural networks*) 18
2.2.4 Redes neurais convolucionais (*convolutional neural networks* - CNNs) 20
2.2.4.1 Camada de convolução (*convolutional layer*) 20
2.2.4.2 Camada de agrupamento (*pooling layer*) 21
2.2.4.3 Camada totalmente conectada (*fully connected layer*) 21
2.2.4.4 Funções de ativação (*activation functions*) 22
2.2.4.5 Funções de perda (*loss functions*) 24
2.2.4.6 Funções de otimização (*optimization functions*) 25
2.2.4.7 Aumentação de dados (*data augmentation*) 25
2.2.4.8 Política de um ciclo (*one cycle policy*) 26
2.2.4.9 Aprendizado por transferência (*transfer learning*) 27
2.2.4.10 Ajuste fino (fine-tuning) 27
2.2.5 Arquiteturas avançadas 28
2.2.5.1 ResNet 28
2.2.5.2 EfficientNet 29
2.2.6 FastAI 29
2.2.7 Explicabilidade 30
2.2.7.1 Mapeamento de ativação de classe ponderado por gradiente (*Gradient-weighted class activation mapping - GradCAM*) 30
2.3 PLACENTAS 31
2.3.1 Anatomia placentária 32
2.3.2 Alterações morfológicas 33
3 ESTADO DA ARTE 36
3.1 METODOLOGIA DE REVISÃO SISTEMÁTICA DE LITERATURA 36
3.1.1 Bancos de dados de pesquisa 36
3.1.2 Termos de busca 36

3.1.3 Procedimento de seleção . . . 36
3.2 TRABALHOS RELACIONADOS . . . 39
3.2.1 PlacentaNet: Automatic Morphological Characterization of Placenta Photos with Deep Learning . . . 39
3.2.2 AI-PLAX: AI-based placental assessment and examination using photos . . . 40
3.2.3 Multi-region saliency-aware learning for cross-domain placenta image segmentation . . . 42
3.2.4 Vision-Language Contrastive Learning Approach to Robust Automatic Placenta Analysis Using Photographic Images . . . 42
3.2.5 Enhancing Automatic Placenta Analysis through Distributional Feature Recomposition in Vision-Language Contrastive Learning . . . 44
4 DESENVOLVIMENTO . . . 45
4.1 CONJUNTO DE IMAGENS . . . 45
4.2 CLASSIFICAÇÃO DE PRESENÇA OU AUSÊNCIA DE ALTERAÇÕES MORFOLÓGICAS . . . 46
4.2.1 Conjunto de dados . . . 46
4.2.2 Redes neurais . . . 48
4.2.3 Hiperparâmetros . . . 48
4.2.3.1 Função de perda . . . 49
4.2.3.2 Otimizador . . . 49
4.2.3.3 Política de um ciclo . . . 49
4.2.3.4 Tamanho do batch . . . 49
4.2.3.5 Número de Épocas e parada antecipada . . . 49
4.2.4 Ambiente de desenvolvimento . . . 49
4.3 CLASSIFICAÇÃO POR TIPO DE ALTERAÇÃO MORFOLÓGICA . . . 49
4.3.1 Conjunto de dados . . . 49
4.3.2 Redes neurais . . . 50
4.3.3 Hiperparâmetros . . . 51
4.3.3.1 Função de perda . . . 51
4.3.3.2 Otimizador . . . 51
4.3.3.3 Política de um ciclo . . . 51
4.3.3.4 Tamanho do batch . . . 52
4.3.3.5 Número de Épocas e parada antecipada . . . 52
4.3.4 Ambiente de desenvolvimento . . . 52
5 RESULTADOS . . . 53
5.1 REDES NEURAIS DE CLASSIFICAÇÃO DE PRESENÇA OU AUSÊNCIA DE ALTERAÇÕES MORFOLÓGICAS EM PLACENTAS . . . 53

5.2 REDES NEURAIS DE CLASSIFICAÇÃO DO TIPO DE ALTERAÇÃO MORFOLÓGICA EM PLACENTAS 57
5.3 DISCUSSÃO DOS RESULTADOS OBTIDOS 63
6 CONCLUSÃO 64
6.1 TRABALHOS FUTUROS 64
REFERÊNCIAS 66
APÊNDICE A – REPOSITÓRIOS 71
APÊNDICE B – ARTIGO 72

1 INTRODUÇÃO

A placenta desempenha papel vital no desenvolvimento fetal, conectando o feto ao útero materno. É responsável por fornecer oxigênio, nutrientes e remover resíduos metabólicos fetais. Adicionalmente, atua como barreira imunológica protetora contra infecções e rejeição materna (LANGSTON *et al.*, 1997; KHONG *et al.*, 2016). Por sua relevância, alterações nesse órgão são indicadores cruciais de desfechos materno-fetais.

Alterações placentárias estão associadas a até 65% das causas de óbito fetal, destacando seu impacto crítico na saúde materno-fetal (ARAÚJO *et al.*, 2006; PTACEK *et al.*, 2014). Assim, quaisquer anomalias na placenta e em suas estruturas relacionadas - como o cordão umbilical e as membranas amnióticas - devem ser identificadas e submetidas a exame anatomopatológico por solicitação ao médico assistente (obstetra) (MITTAL *et al.*, 2020; DEL DUQUE *et al.*, 2022). Entretanto, ainda que o exame forneça informações confiáveis sobre o desenvolvimento neonatal e planejamento familiar para futuras gestações (KHONG *et al.*, 2016; CHANG, 2009), sua realização para todos os partos sem intercorrências é onerosa e inviável. Dessa forma, sua realização só ocorre após análise macroscópica do órgão e definição de necessidade pelo médico responsável.

Diante dessas limitações, o uso de inteligência artificial tem emergido como uma solução promissora na análise de imagens médicas, especialmente com o emprego de técnicas de deep learning (LITJENS *et al.*, 2017). Embora avanços tenham sido feitos na análise de imagens pré-parto, como ultrassonografias e ressonâncias magnéticas (ALANSARY *et al.*, 2016; MALATHI; SHANTHI, 2011; LOONEY *et al.*, 2017), sua aplicação no exame macroscópico pós-parto permanece escassa (PAN *et al.*, 2023).

As Redes Neurais Convolucionais (CNNs) são uma das principais ferramentas de inteligência artificial para análise de imagens. Compostas por camadas especializadas em identificar padrões complexos, essas redes têm revolucionado aplicações médicas (LECUN; BENGIO; HINTON, 2015). Por conta disso, sua eficácia será explorada neste estudo.

Atualmente existem redes neurais pré-treinadas - ResNet, VGG, Inception, EfficientNet e YOLO - que são utilizadas como uma base poderosa ao desenvolvimento de modelos específicos. Para isso, utiliza-se a técnica de aprendizado por transferência (*transfer learning*) para a adaptação delas, sendo então capazes de solucionar diferentes tarefas para as quais foram concebidas. No geral, esse método é responsável por garantir uma melhora na performance em relação a modelos treinados do zero (TAJBAKHSH *et al.*, 2016).

Em vista disso, a presente tese visa desenvolver e avaliar a eficácia de redes neurais convolucionais desenvolvidas a partir de modelos pré-treinados na caracterização macroscópica placentária, verificando possíveis alterações preditivas de doenças. Dessa forma, pode-se utilizar os modelos para otimizar o processo de diagnóstico e proporcionar melhorias ao cuidado materno-fetal.

1.1 JUSTIFICATIVA

O exame anatomopatológico da placenta desempenha um papel crucial na investigação de desfechos materno-fetais desfavoráveis, como prematuridade, restrição de crescimento intrauterino e perdas gestacionais. Ele possibilita a identificação de alterações com implicações diretas no manejo de futuras gestações e na saúde materna.

No entanto, considerando que a maioria dos partos ocorre sem intercorrências, a realização rotineira do exame anatomopatológico em todos os casos mostra-se inviável do ponto de vista operacional, logístico e financeiro, além de frequentemente desnecessária. Por essa razão, a seleção dos casos passíveis de investigação depende de uma análise macroscópica criteriosa realizada por um especialista, a qual, embora seja um procedimento consolidado, demanda tempo, está sujeita à variabilidade interobservador e pode ter sua qualidade comprometida pelo elevado fluxo de atendimentos em hospitais e clínicas.

Diante desse cenário, a presente tese propõe o desenvolvimento de um sistema baseado em aprendizado profundo para reconhecimento automatizado de alterações macroscópicas na placenta. Essa abordagem visa otimizar o processo de solicitação de exame anatomopatológico ao proporcionar maior precisão na triagem, reduzindo a realização de exames desnecessários e, consequentemente, os custos associados, ao mesmo tempo em que aumenta a probabilidade de detecção de condições clinicamente relevantes.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Desenvolver redes neurais convolucionais baseadas em modelos pré-treinados para a classificação de alterações preditivas de doenças placentárias, garantindo uma acurácia mínima de 90% nos modelos propostos.

1.2.2 Objetivos Específicos

- Identificar e classificar a presença ou ausência de anomalias de forma, coloração e vascularização nas placentas, baseando-se em critérios médicos padronizados.
- Identificar e classificar os tipos de anomalias de forma, coloração e vascularização nas placentas, baseando-se em critérios médicos padronizados.
- Validar os modelos desenvolvidos por meio de testes com imagens de placentas inéditas, utilizando métricas como acurácia, precisão, *recall* e *F1-Score*.

2 CONCEITOS RELACIONADOS

Neste capítulo são apresentados os principais conceitos para a realização deste trabalho. Este trabalho fundamenta-se em princípios da visão computacional, da inteligência artificial e da medicina.

2.1 VISÃO COMPUTACIONAL

Visão computacional é um subcampo da ciência da computação dedicado a estabelecer métodos para extrair, interpretar e analisar informações significativas a partir de dados visuais, como imagens e vídeos. Historicamente a visão computacional passou por evoluções de paradigma, saindo de uma abordagem algorítmica para a utilização de inteligência artificial, conforme descrito na seção 2.2. Dentre as variadas aplicações da visão computacional, como a detecção de objetos, a segmentação semântica e a classificação de imagens, destaca-se a última no presente trabalho. Essa abordagem será explorada na seção 2.1.1.

2.1.1 Classificação de imagens

A classificação de imagens é uma das tarefas mais fundamentais da visão computacional. Consiste em atribuir um rótulo ou categoria a uma imagem com base em suas características visuais. Essa tarefa é amplamente utilizada em diversos domínios, incluindo diagnóstico médico, reconhecimento facial, inspeção de qualidade em fábricas, entre outros (WANG; ZHU, 2023).

Existem diferentes abordagens para a classificação de imagens, que podem ser divididas em:

- **Não supervisionada:** consiste em agrupar imagens com base em similaridades visuais sem o uso de rótulos prévios. Algoritmos de *clustering*, como *k-means* ou agrupamento hierárquico, são frequentemente utilizados. Essa abordagem é útil quando não há dados anotados disponíveis.

- **Supervisionada:** consiste em treinar um modelo com um conjunto de imagens previamente rotuladas (*labeled data*), de modo que ele aprenda a associar padrões visuais às respectivas classes. Cada exemplo no conjunto de treinamento consiste em um objeto de entrada - denominado *feature* - e um valor de saída desejado - denominado *target*. O objetivo desse tipo de aprendizado é que o modelo se torne capaz de generalizar a relação entre entradas e saídas, de forma a ser capaz de prever corretamente o rótulo para dados nunca antes vistos.

2.2 INTELIGÊNCIA ARTIFICIAL (IA)

A inteligência artificial é um ramo da ciência da computação que compreende um conjunto de tecnologias concebidas para capacitar sistemas computacionais a executar tarefas que, historicamente, demandam inteligência humana. Tais tarefas incluem capacidades avançadas como a percepção visual, a compreensão e a tradução de linguagens, a análise de dados complexos e a tomada de decisões ou formulação de recomendações.

Essa tecnologia possui diversas ramificações, como:

- Aprendizado de máquina
- Sistemas especialistas
- Processamento de linguagem natural (*natural language processing* - NLP)
- Visão computacional

2.2.1 Aprendizado de máquina (*machine learning* - ML)

O aprendizado de máquina é um subcampo da inteligência artificial que se concentra no desenvolvimento de algoritmos que permitem aos sistemas computacionais aprender padrões a partir de dados previamente obtidos, sem serem explicitamente programados para cada tarefa específica. Dessa forma, em vez de necessitar da codificação de regras específicas, um modelo de ML ajusta parâmetros internos com base nos dados aos quais é exposto e melhora seu desempenho em determinada tarefa iterativamente.

2.2.2 Aprendizado profundo (*deep learning* - DL)

O aprendizado profundo é um subconjunto especializado do ML que se baseia em redes neurais artificiais com múltiplas camadas, conhecidas como redes neurais profundas. O termo profundo refere-se à presença de um número significativo de camadas ocultas entre a camada de entrada e a de saída.

A principal vantagem do DL é a capacidade de aprender representações hierárquicas de características. Em uma tarefa de visão computacional, por exemplo, as camadas iniciais de uma rede profunda aprendem a detectar características simples e de baixo nível - bordas, cantos e gradientes de cor. As camadas subsequentes aprendem a combinar essas - características simples para reconhecer padrões mais complexos e abstratos - texturas, formas e objetos. Essa capacidade elimina a necessidade de engenharia de características manual.

2.2.3 Redes neurais artificiais (*artificial neural networks*)

Uma rede neural artificial é um modelo computacional inspirado na estrutura e no funcionamento do cérebro humano. Consiste em um conjunto de unidades de processa-

mento interconectadas, denominadas neurônios ou nós. Estes nós estão organizados em camadas sequenciais:

- Camada de entrada (*input layer*): recebe os dados brutos
- Camada(s) oculta(s) (*hidden layer(s)*): responsável(is) pela extração de padrões
- Camada de saída (*output layer*): produz o resultado final da tarefa, como uma predição de classe

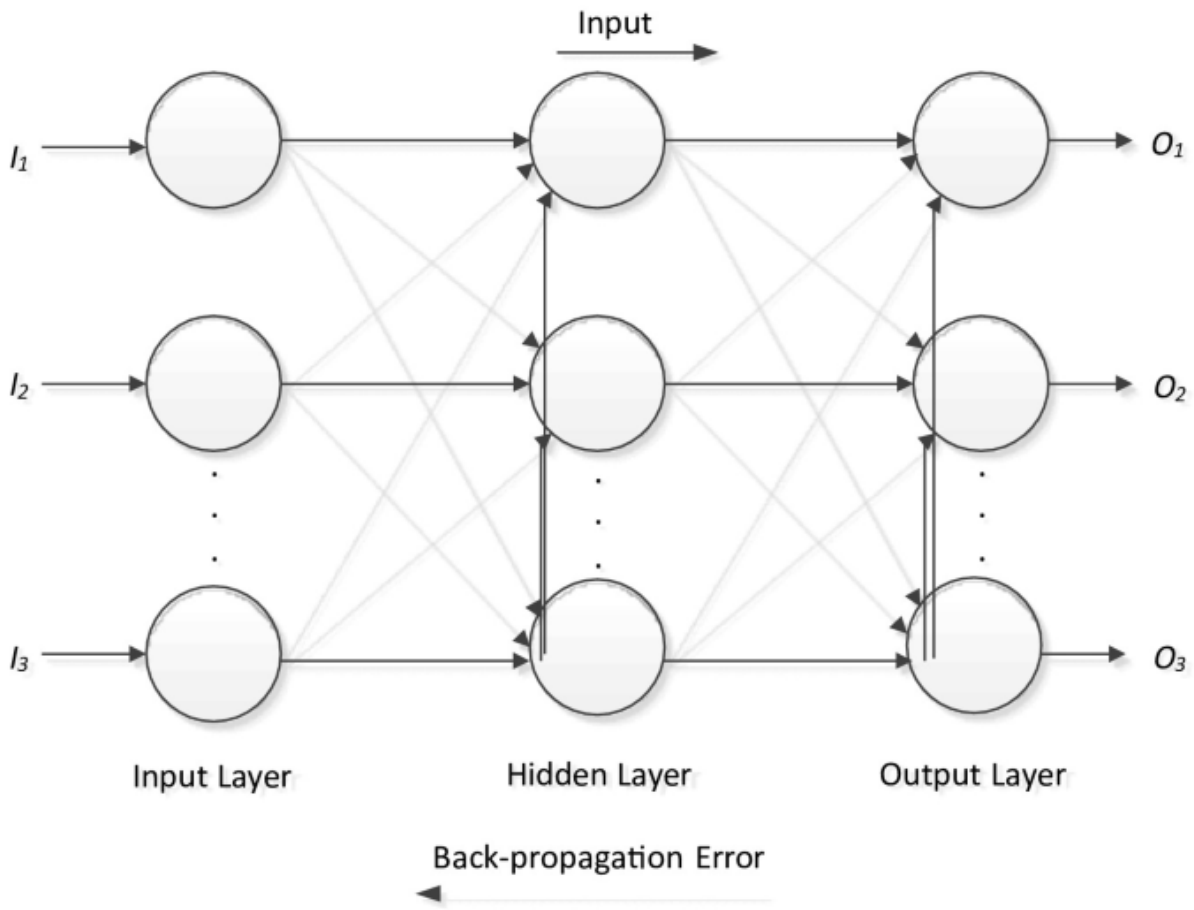


Figura 1 – Um esquema representativo de uma rede neural artificial (DAS; SADIQ; MIR-JALILI, 2025).

Cada conexão entre neurônios de camadas adjacentes possui um peso associado e um *bias*, de modo a modular a força do sinal de entrada e ajustar a saída independentemente das entradas. Durante o processo de treinamento, a rede modifica iterativamente os valores para os pesos e *biases* de modo a se adequar ao conjunto de entradas e saídas observado.

Um dos problemas das redes neurais tradicionais é o fato de serem comumente estruturadas com camadas totalmente conectadas, onde cada neurônio está conectado a todos os neurônios da camada seguinte. Para o processamento de imagens, essa arquitetura exige que a matriz de pixels 2D seja transformada em um vetor unidimensional. Esse processo causa uma perda de informação espacial, já que pixels vizinhos são tratados da

mesma forma que pixels distantes. Além disso, essa abordagem leva a uma explosão no número de parâmetros, o que dificulta o custo de treinamento (LECUN; BOTTOU *et al.*, 1998). Para resolver esse problema, foram desenvolvidas as redes neurais convolucionais, expostas na seção 2.2.4.

2.2.4 Redes neurais convolucionais (*convolutional neural networks* - CNNs)

Redes neurais convolucionais são redes neurais especializadas no processamento de dados com estrutura espacial, como imagens. Elas foram projetadas para superar as limitações das redes tradicionais ao lidar com dados visuais, preservando a relação espacial entre os pixels e reduzindo significativamente o número de parâmetros do modelo.

Inspiradas no funcionamento do córtex visual dos animais, as CNNs utilizam filtros (ou *kernels*) que percorrem a imagem em pequenas regiões locais, identificando padrões como bordas, texturas e formas. Esses filtros são aprendidos automaticamente durante o treinamento e permitem que a rede extraia características relevantes de forma hierárquica, desde os detalhes mais simples até os mais complexos.

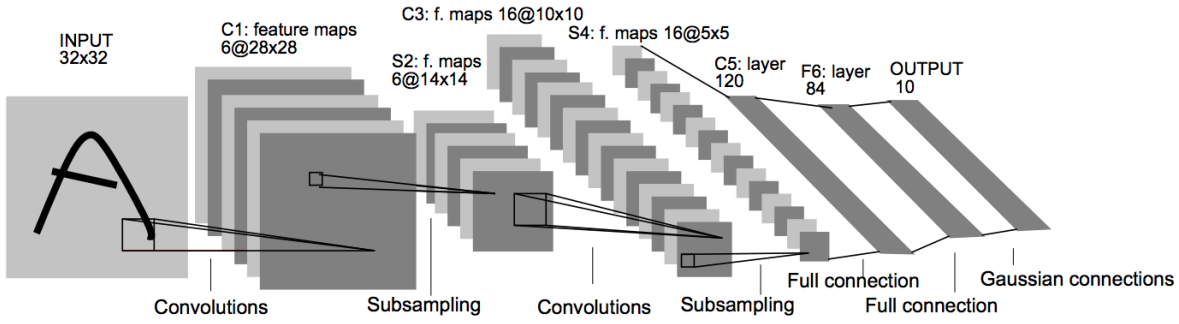


Figura 2 – Um esquema representativo de uma rede neural convolucional (LECUN; BOTTOU *et al.*, 1998).

2.2.4.1 Camada de convolução (*convolutional layer*)

A camada de convolução é o bloco de construção fundamental das CNNs. Sua operação central é a convolução matemática, que consiste em aplicar um ou mais filtros (*kernels*) sobre a imagem de entrada. *Kernels* são pequenas matrizes de pesos que atuam como um descritor de características. Em cada posição, é realizada uma operação de produto escalar - multiplicação elemento a elemento seguida de uma soma - entre os valores do *kernel* e a porção da imagem sob ele, de modo que a aplicação do *kernel* para todas as posições produza um mapa de ativação (*feature map*) que sirva para extrair padrões locais - bordas, cores, formas, texturas, entre outros.

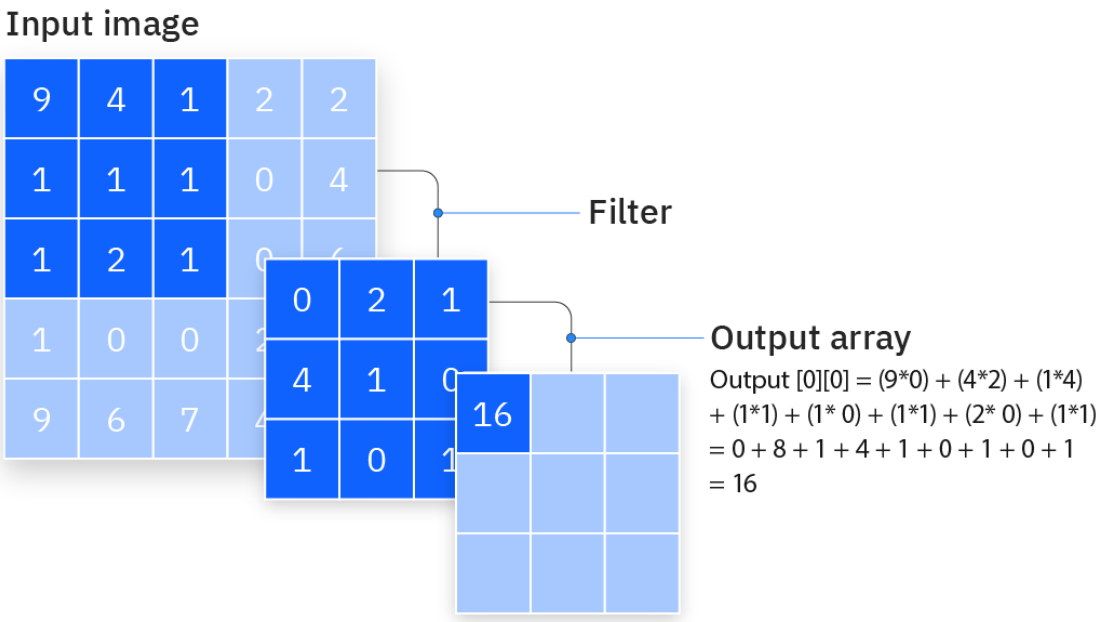


Figura 3 – Um esquema representativo de uma camada de convolução (IBM, 2025).

2.2.4.2 Camada de agrupamento (*pooling layer*)

A camada de *pooling* tem como objetivo reduzir a dimensionalidade espacial (largura e altura) dos mapas de ativação gerados pelas camadas de convolução, preservando apenas as informações mais relevantes. Essa redução contribui para a diminuição do custo computacional e ajuda a tornar o modelo mais robusto a pequenas variações ou deslocamentos na imagem de entrada, auxiliando a prevenir o *overfitting*.

Os métodos de *pooling* mais comuns são:

- ***Max Pooling*:** seleciona o maior valor dentro de uma janela da matriz de ativação.
- ***Average Pooling*:** calcula a média dos valores dentro da janela.
- ***L2 Pooling*:** computa a raiz quadrada da soma dos quadrados dos valores na janela.
- ***Overlapping Pooling*:** aplica *pooling* com janelas que se sobrepõem, aumentando a resolução dos mapas de ativação.
- ***Spatial Pyramid Pooling*:** realiza *pooling* em múltiplas escalas espaciais e concatena os resultados, permitindo entrada de tamanho variável sem distorcer a estrutura espacial.

2.2.4.3 Camada totalmente conectada (*fully connected layer*)

A camada totalmente conectada é normalmente utilizada nas últimas etapas de uma rede neural convolucional como responsável por tomar a decisão final do modelo,

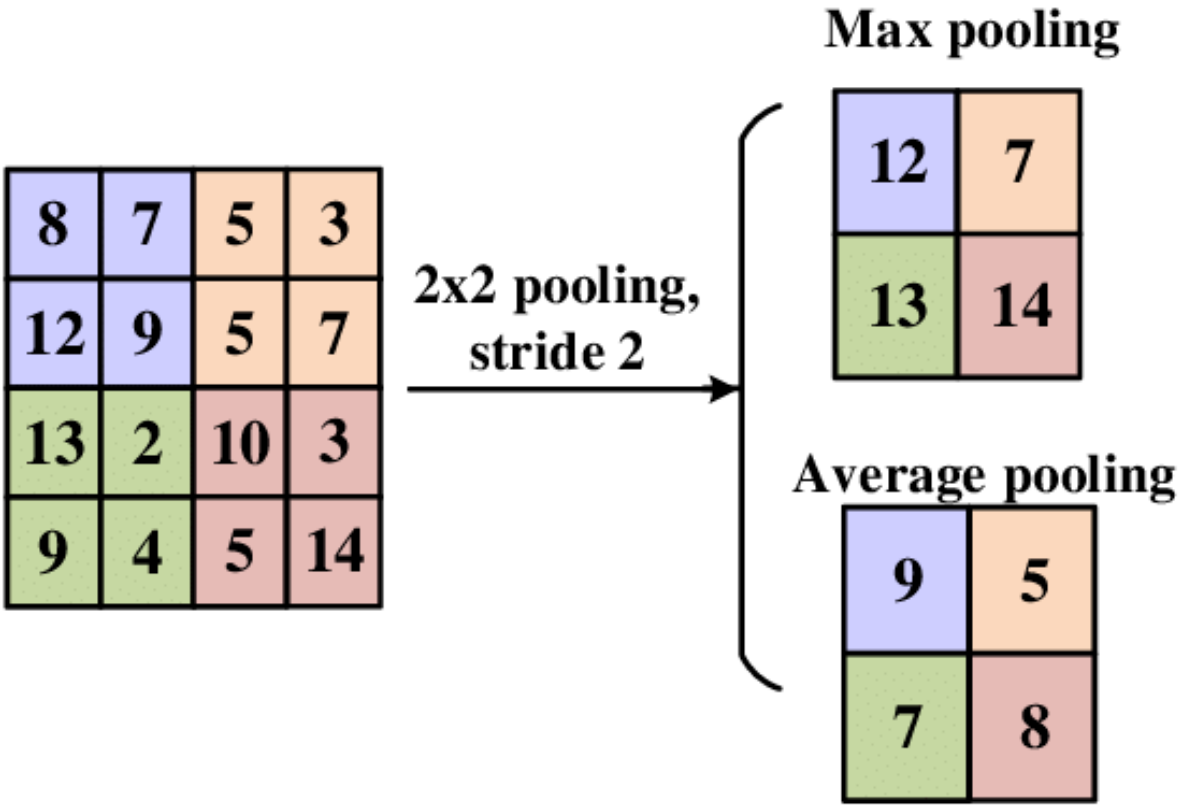


Figura 4 – Um esquema do funcionamento de uma *pooling layer* (YINGGE; ALI; LEE, 2020).

como classificar uma imagem entre diferentes categorias. Nessa camada, cada neurônio é conectado a todos os neurônios da camada anterior, permitindo a combinação global das características extraídas nas fases anteriores da rede. Além disso, os mapas de ativação são transformados em um único e longo vetor unidimensional em um processo conhecido como *flattening*.

2.2.4.4 Funções de ativação (*activation functions*)

As funções de ativação são operações aplicadas após a realização de operações lineares - convolução, multiplicação por pesos, entre outros - como forma de introduzir não linearidade ao modelo. Com isso, o modelo é capaz de aprender padrões complexos e não lineares.

- ***Sigmoid*:** definida como $f(x) = \frac{1}{1+e^{-x}}$, comprime a entrada em um intervalo entre 0 e 1. Tende a sofrer com o gradiente devanscente, principalmente em redes profundas.
- **Tanh (tangente hiperbólica):** definida como $f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$, mapeia os valores para o intervalo entre -1 e 1. Sofre com o gradiente pequeno em entradas saturadas, porém obtém melhor desempenho ao centralizar a saída em torno de zero.

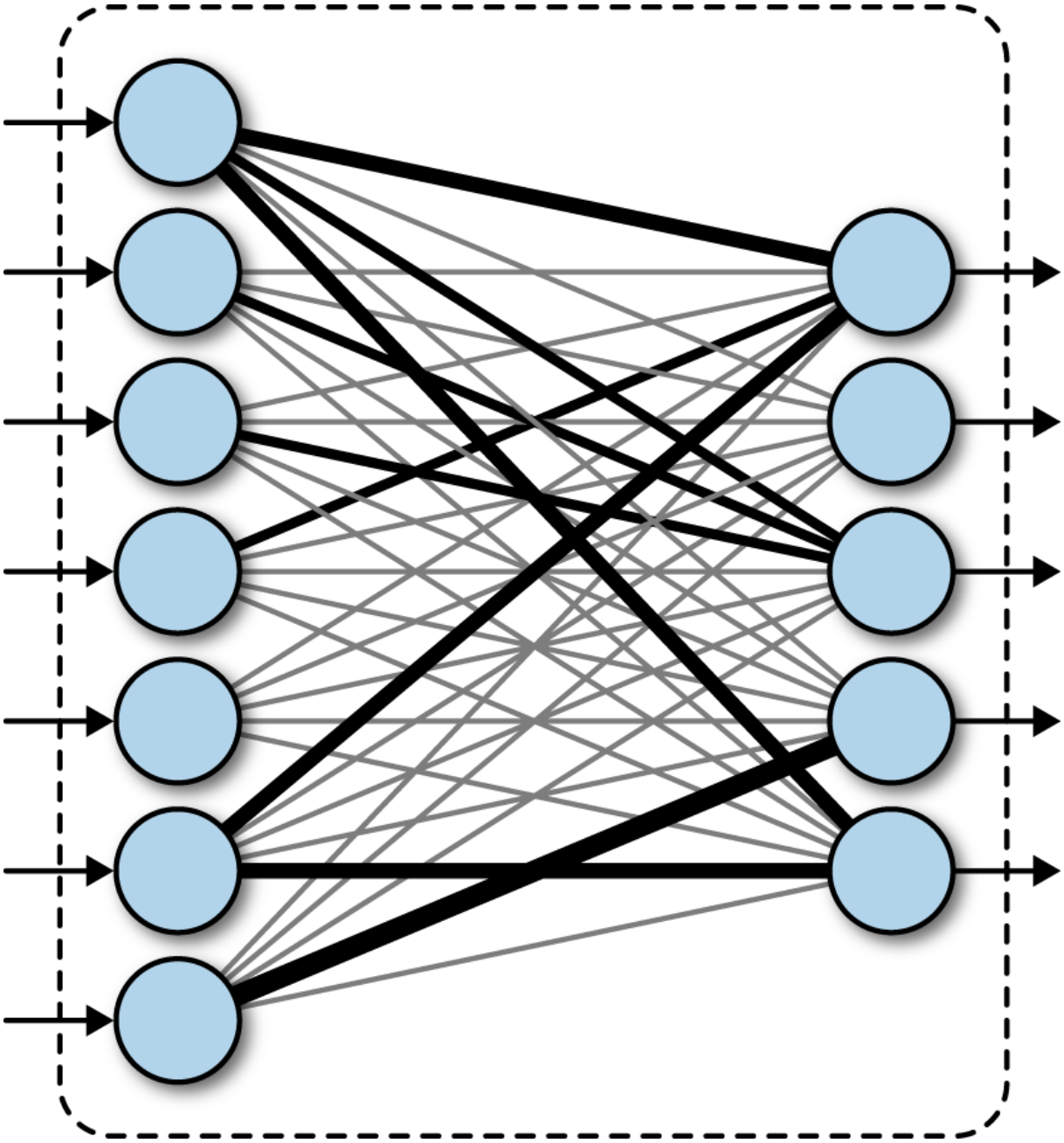


Figura 5 – Um esquema representativo de uma *fully connected layer* (HOPE; RESHEFF; LIEDER, 2018).

- **ReLU (*Rectified Linear Unit*):** definida como $f(x) = \max(0, x)$, essencialmente substitui os valores negativos do mapa de ativação por zero. Dessa forma, é capaz de mitigar o problema do gradiente devanescente, além de ser computacionalmente eficiente.

- ***Leaky ReLU*:** variante da ReLU que mantém valores negativos com um pequeno coeficiente, definida como $f(x) = \max(\alpha x, x)$, onde α é um valor pequeno (geralmente 0,01). Essa função busca contornar o problema dos neurônios mortos, permitindo que unidades com ativações negativas ainda aprendam.

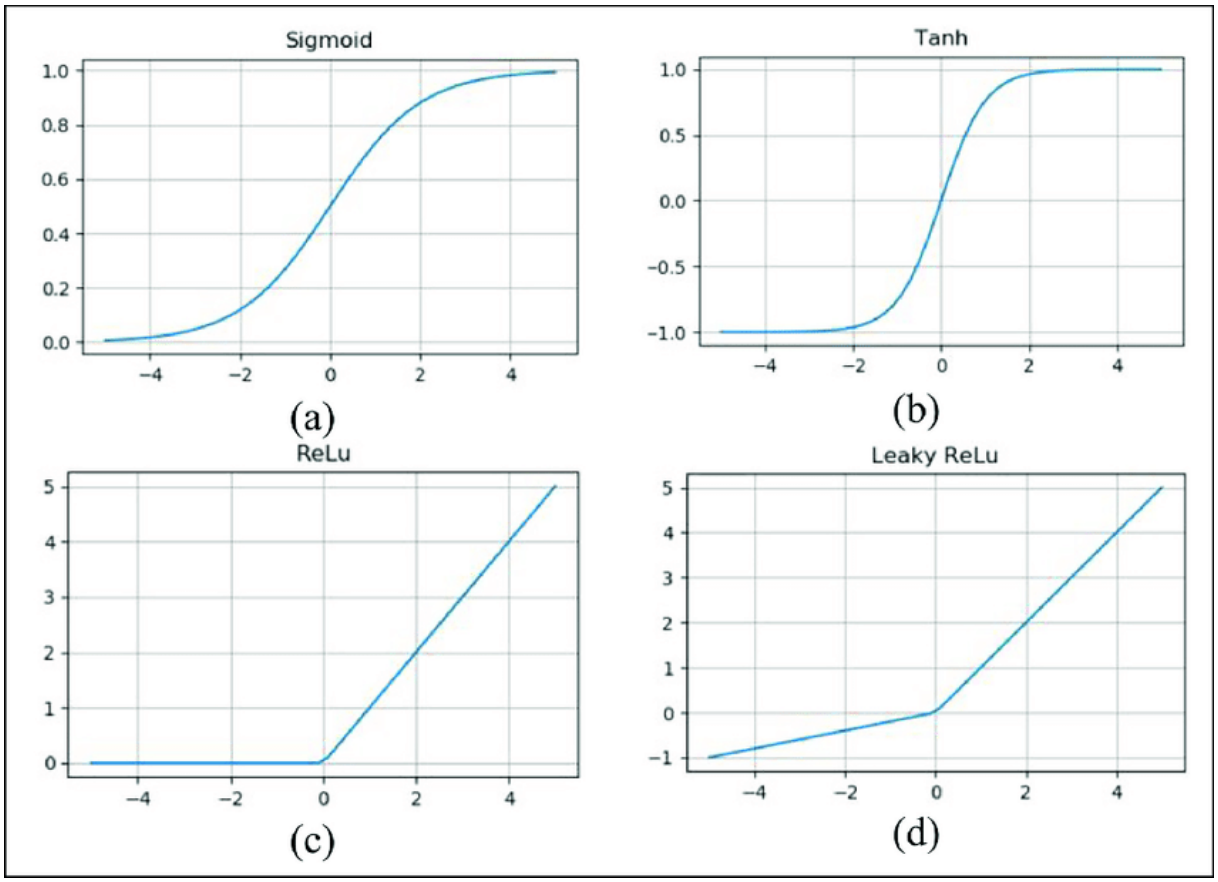


Figura 6 – Um comparativo das funções de ativação sigmoid, tanh, ReLU e Leaky ReLU (KUO; LIN; HU, 2020).

2.2.4.5 Funções de perda (*loss functions*)

A função de perda quantifica a diferença entre a predição da rede e o valor real esperado. Ela fornece um sinal de erro que guia o processo de aprendizado, permitindo que a rede ajuste seus parâmetros para melhorar o desempenho. Possuem diversos tipos:

- *Binary Cross-Entropy:* usada em classificações binárias.

- *Categorical Cross-Entropy:* usada em classificações multiclasse.

- *Mean Squared Error (MSE):* utilizada principalmente em regressão.
- *Focal Loss:* útil em cenários com classes desbalanceadas.

2.2.4.6 Funções de otimização (*optimization functions*)

As funções de otimização são responsáveis por atualizar os parâmetros da rede neural durante o treinamento, com o objetivo de minimizar a função de perda. Utilizam os gradientes calculados por meio do algoritmo de retropropagação (*backpropagation*) para guiar a rede em direção a soluções com melhor desempenho.

A escolha do otimizador pode influenciar significativamente a velocidade de convergência, a estabilidade do treinamento e a qualidade da solução final. A seguir, são apresentados os algoritmos de otimização mais utilizados em redes neurais convolucionais:

- ***Gradient Descent*:** algoritmo de otimização que atualiza os parâmetros da rede na direção do gradiente negativo da função de perda, com o objetivo de minimizá-la. Requer o cálculo do gradiente sobre todo o conjunto de dados, o que pode ser computacionalmente custoso em grandes volumes.
- **SGD (*Stochastic Gradient Descent*):** é a versão estocástica do método clássico de gradiente descendente. Atualiza os parâmetros com base em pequenos lotes (mini-batches) de dados, tornando o treinamento mais rápido e escalável.
- **RMSProp (*Root Mean Square Propagation*):** adapta a taxa de aprendizado para cada parâmetro com base na média móvel dos quadrados dos gradientes.
- **Adam (*Adaptive Moment Estimation*):** combina as vantagens do SGD com momentum e do RMSProp. Utiliza médias móveis dos gradientes e dos quadrados dos gradientes para adaptar individualmente a taxa de aprendizado de cada parâmetro (KINGMA; BA, 2015).

2.2.4.7 Aumentação de dados (*data augmentation*)

A aumentação de dados é uma técnica amplamente utilizada no treinamento de redes neurais para aumentar a diversidade e a quantidade efetiva de exemplos no conjunto de dados, sem a necessidade de coletar novas imagens. Isso é feito por meio da aplicação de transformações sobre as imagens originais, preservando suas características essenciais, mas introduzindo variações que ajudam o modelo a generalizar melhor.

Essa técnica é especialmente importante em contextos com conjuntos de dados limitados ou desbalanceados. Além disso, ela atua como uma forma de regularização, contribuindo para a redução do *overfitting*.

As transformações aplicadas podem ser:

- **Geométricas:** rotação, espelhamento horizontal ou vertical, translação, escalonamento (zoom), recorte (*cropping*), cisalhamento (*shear*).
- **Fotométricas:** ajuste de brilho, contraste, saturação, adição de ruído, variações de cor.
- **Específicas para visão computacional:** *cutout* (remoção de regiões aleatórias da imagem), *mixup* (combinação de duas imagens) e *random erasing* (substituição dos pixels de uma seção da imagem por valores aleatórios).

As operações de aumentação podem ser aplicadas de forma aleatória durante o treinamento, resultando em um conjunto virtualmente ilimitado de amostras distintas, o que aumenta a robustez do modelo e melhora sua capacidade de generalização.

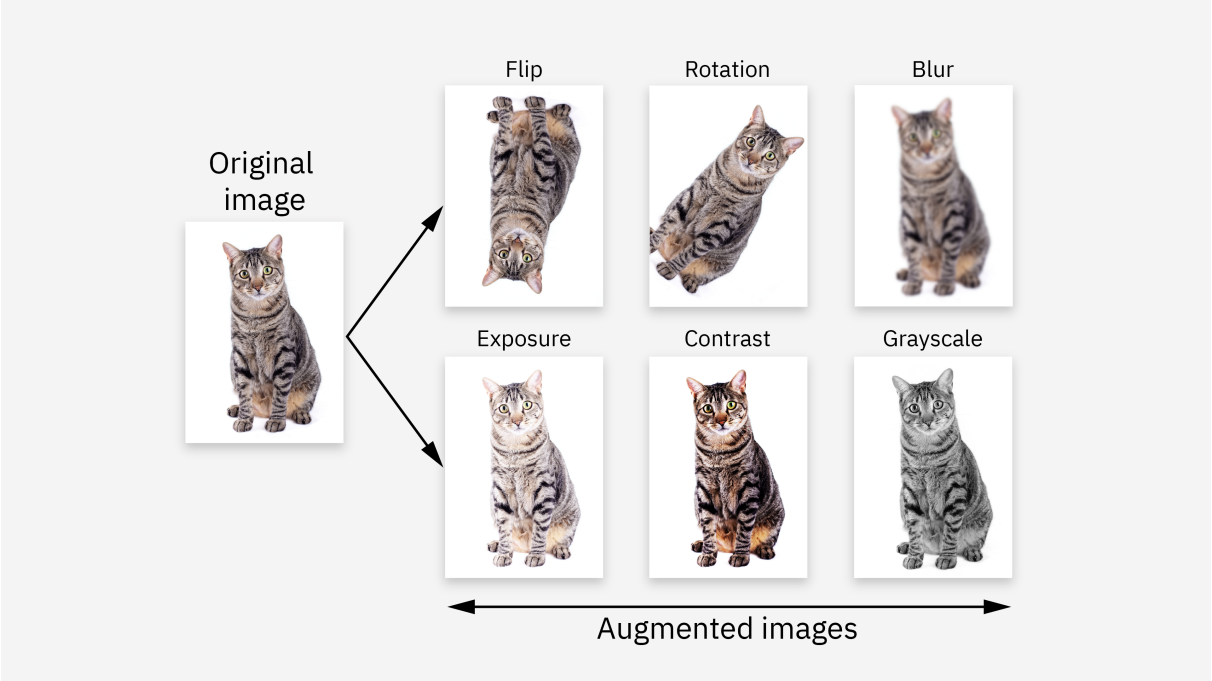


Figura 7 – Exemplos de aumentação de dados (MUREL; KAVLAKOGLU, 2024).

2.2.4.8 Política de um ciclo (*one cycle policy*)

A *One Cycle Policy* (SMITH; TOPIN, 2017) é uma estratégia de ajuste dinâmico da taxa de aprendizado e do momentum durante o treinamento de redes neurais. Essa política visa acelerar a convergência e melhorar a capacidade de generalização do modelo.

O princípio fundamental da técnica consiste em variar a taxa de aprendizado ao longo do tempo de forma cíclica: primeiro aumentando gradualmente até um valor máximo e, em seguida, diminuindo até um valor muito inferior ao inicial. Simultaneamente, o momentum é ajustado de forma inversa, diminuindo quando a taxa de aprendizado aumenta e subindo quando esta é reduzida. Esse comportamento força a rede a explorar

regiões mais amplas do espaço de parâmetros no início e se estabilizar próximo a um mínimo global ao final do treinamento.

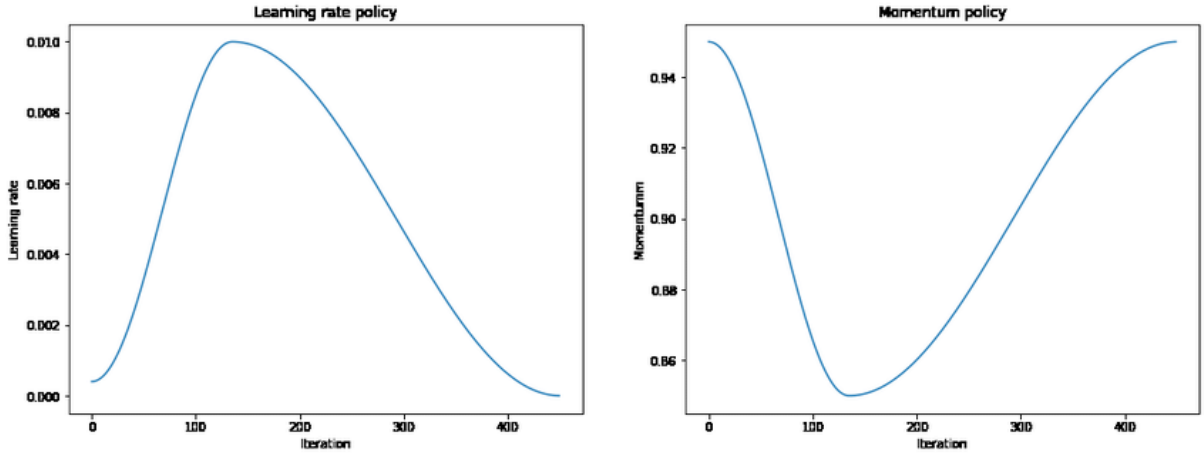


Figura 8 – Ilustração da variação da taxa de aprendizado e do momentum com a *one cycle policy* (ROMERO *et al.*, 2019).

2.2.4.9 Aprendizado por transferência (*transfer learning*)

O *transfer learning* é uma estratégia eficaz para contornar o problema de escassez de dados e para reduzir o custo computacional para o treinamento de redes neurais. Essa técnica consiste em aproveitar o conhecimento adquirido por um modelo pré-treinado em um dataset de grande escala. A lógica é que as camadas iniciais de uma CNN aprendam a reconhecer características de baixo nível, como bordas, texturas, cores e formas. Com isso, redes específicas podem partir de um estágio inicial avançado, com pesos otimizados para o reconhecimento das características universais.

O método de treinamento padrão com a técnica de transfer learning é optar pelo congelamento dos pesos iniciais, treinando apenas a parte final da rede para a tarefa de classificação, segmentação ou detecção específica ao problema. Pode-se então optar pela realização de um ajuste fino, conforme discutido na seção 2.2.4.10.

2.2.4.10 Ajuste fino (fine-tuning)

O *fine-tuning* é uma etapa complementar ao aprendizado por transferência, na qual se realiza o ajuste dos pesos de camadas adicionais - ou até mesmo de toda a rede - após o treinamento inicial com os pesos congelados. Essa abordagem permite que o modelo pré-treinado se adapte de forma mais específica ao novo conjunto de dados, refinando os filtros previamente aprendidos com base nas particularidades da nova tarefa. Essa técnica é especialmente útil quando o novo domínio de aplicação difere consideravelmente daquele em que o modelo original foi treinado.

2.2.5 Arquiteturas avançadas

2.2.5.1 ResNet

Uma premissa ao desenvolvimento de redes neurais é que modelos mais profundos, com mais camadas, deveriam ter capacidades de representação maiores devido à capacidade de generalização crescente ao longo do modelo. Consequentemente, espera-se que esses modelos alcancem resultadores melhores. No entanto, experimentos práticos revelaram um fenômeno contraintuitivo de degradação de desempenho para grandes modelos. Observou-se que a partir de certo ponto, adicionar camadas à rede piorava os resultados de treinamento e teste (HE *et al.*, 2016).

Esse problema de degradação está intrinsicamente ligado a um desafio conhecido como o problema do gradiente devanescente (*vanishing gradient*). Durante o processo de *backpropagation*, o gradiente da função de perda é propagado através da rede da camada de saída à de entrada. Caso a rede utilize funções de ativação cujas derivadas possuem valor máximo inferior a 1, o produto contínuo desses pequenos valores diminui o gradiente exponencialmente. Com isso, camadas mais próximas da entrada apresentam pequena atualização no treinamento, o que estagna o processo de treinamento.

Como resposta a esse problema, foi introduzido o modelo ResNet (HE *et al.*, 2016), permitindo o treinamento de redes substancialmente mais profundas que as anteriores. A ideia central do modelo é o bloco residual (*residual block*). Esse bloco consiste de duas ou três camadas convolucionais e uma conexão de atalho (*skip connection* ou *shortcut connection*). O papel da *skip connection* é adicionar a entrada do bloco à saída das camadas convolucionais. Com isso, surge um caminho alternativo ao fluxo de informação e para o processo de *backpropagation*, o que garante com que o gradiente total não desapareça, mesmo que se torne pequeno. Essa inovação permitiu a otimização de redes com mais de 50 camadas, as quais superaram arquiteturas mais rasas em tarefas de classificação.

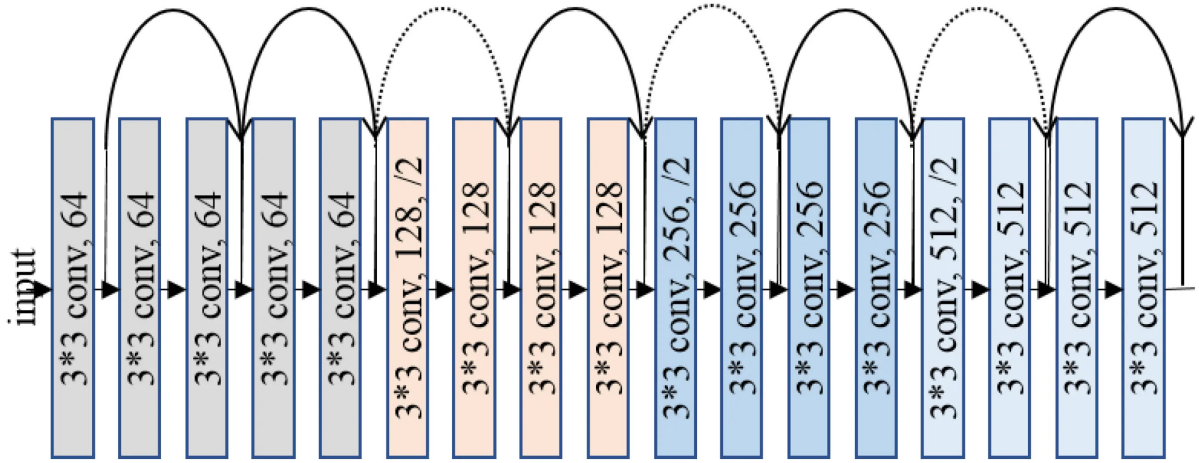


Figura 9 – Representação gráfica do modelo ResNet (HE *et al.*, 2016).

2.2.5.2 EfficientNet

A arquitetura EfficientNet (TAN; LE, 2019) introduziu uma abordagem fundamentalmente nova para o escalonamento de modelos, denominada escalonamento composto (*compound scaling*). A contribuição da arquitetura foi de que as três dimensões do escalonamento - profundidade, largura e resolução - devem ser balanceadas de forma coordenada, ao invés de escaladas separadamente. O escalonamento composto propõe realizar um escalonamento com um coeficiente composto aplicado uniformemente à profundidade, largura e resolução da rede.

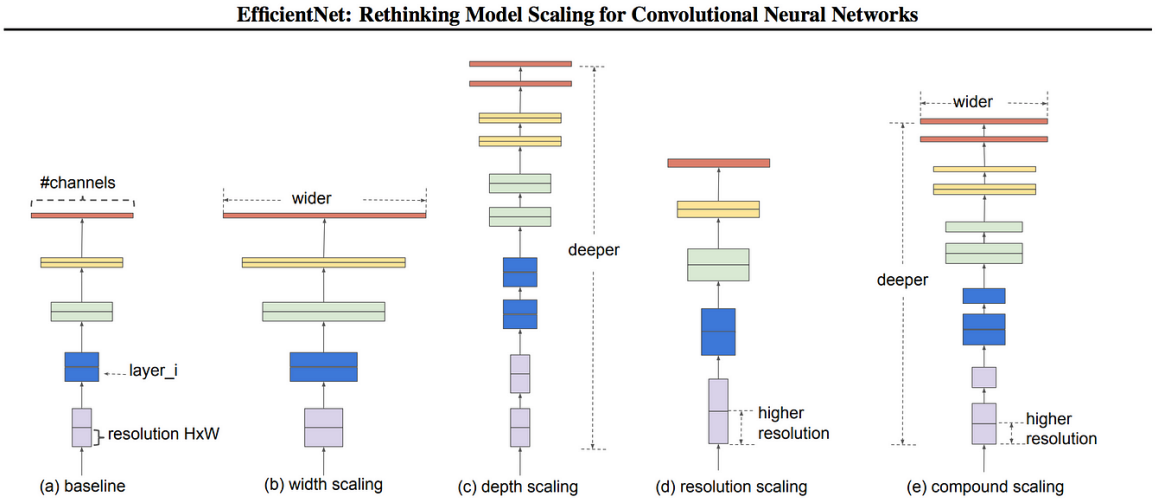


Figura 10 – Representação gráfica dos conceitos de profundidade, largura e resolução para o escalonamento composto (TAN; LE, 2019).

Para aplicar o princípio, os autores propuseram uma arquitetura base chamada EfficientNet-B0, em que o principal bloco de construção é o MBConv (*Mobile Inverted Bottleneck Convolution*). Este bloco divide uma convolução padrão nas etapas de convolução em profundidade (*depthwise convolution*) - que aplica um filtro a cada canal de entrada de forma independente - e convolução pontual (*pointwise convolution*) - que combina as saídas da etapa anterior. Aliado a isso, a arquitetura incorpora módulos *Squeeze-and-Excitation* (SE). O módulo "aperta"(*squeezes*) a informação espacial de cada canal em um único descritor e, em seguida, usa uma pequena rede neural para gerar pesos que "excita"(*excites*) ou amplifica os canais mais informativos, enquanto suprime os menos relevantes.

2.2.6 FastAI

O *FastAI* (HOWARD; GUGGER, 2020) é um framework de aprendizado profundo de alto nível, desenvolvido com o objetivo de tornar o uso de técnicas avançadas de inteligência artificial mais acessível, eficiente e reprodutível. Ele é construído sobre o *PyTorch*, uma das bibliotecas para aprendizado profundo, oferecendo uma interface mais intuitiva e de alto nível para tarefas comuns de *deep learning*. A biblioteca abstrai diversas

etapas complexas do processo de treinamento de modelos, como a preparação de dados, escolha de arquiteturas, definição de métricas e ajustes de hiperparâmetros.

2.2.7 Explicabilidade

A Explicabilidade em Inteligência Artificial (*XAI - explainable AI*) é um campo emergente que visa focar no desenvolvimento de técnicas que tornam as decisões dos modelos de ML mais compreensíveis para os seres humanos. As técnicas de XAI podem ser categorizadas de diversas maneiras, incluindo o escopo de explicabilidade (global ou local), o momento de sua aplicação (pré ou pós-hoc) e a dependência do modelo (agnósticas ou específicas) (GUIDOTTI *et al.*, 2018).

Dentre as técnicas mais relevantes para o contexto de visão computacional e modelos baseados em redes neurais, destacam-se:

- **LIME (*local interpretable model-agnostic explanations*):** Técnica pós-hoc e agnóstica que cria um modelo substituto local e interpretável para explicar as predições de instâncias individuais, mostrando quais características de entrada (pixels ou superpixels) são mais relevantes para o resultado (RIBEIRO; SINGH; GUESTRIN, 2016).

- **SHAP (*Shapley additive explanations*):** Baseado na teoria dos jogos de Shapley, atribui a importância de cada característica (pixel, superpixel ou característica do modelo) para a predição, garantindo que a soma das contribuições seja igual à diferença entre a predição e o valor médio das predições (LUNDBERG; LEE, 2017).

- **Mapas de saliência (*saliency maps*):** Determinam a relevância de cada pixel de entrada para a classificação, geralmente calculando o gradiente da saída da rede em relação à entrada. O resultado é um mapa de calor onde pixels mais brilhantes (ou de cor mais intensa) são mais importantes (SIMONYAN; VEDALDI; ZISSERMAN, 2014).

- **Mapeamento de ativação de classe (*class activation mapping - CAM*):** Uma técnica específica para modelos com camadas convolucionais seguidas por uma *Global Average Pooling* (GAP) e uma camada totalmente conectada, que permite visualizar as regiões da imagem mais relevantes para uma determinada classe (ZHOU *et al.*, 2016)

2.2.7.1 Mapeamento de ativação de classe ponderado por gradiente (*Gradient-weighted class activation mapping - GradCAM*)

O Grad-CAM (SELVARAJU *et al.*, 2017) é uma evolução do CAM que se tornou uma das técnicas de XAI mais amplamente utilizadas, pois elimina a necessidade de

modificação da arquitetura da rede, podendo ser aplicada a virtualmente qualquer CNN para gerar mapas de calor visuais de alta resolução.

O Grad-CAM foca em identificar as regiões da imagem que influenciam a predição de uma classe específica. Ele realiza um cálculo com a importância de cada mapa de características na última camada convolucional da rede em relação à pontuação da classe de interesse, como evidenciam os cálculos abaixo.

1. **Cálculo do gradiente:** O gradiente da pontuação da classe c (y^c) em relação aos mapas de características A^k da última camada convolucional é calculado:
$$\frac{\partial y^c}{\partial A^k}$$
2. **Cálculo do peso do neurônio (α_k^c):** O gradiente é então submetido a uma Média Global (*Global Average Pooling* - GAP) para obter os pesos α_k^c. Este peso representa a importância do mapa de características k para a classe c:
$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}$$
3. **Geração do mapa de calor:** O mapa de calor ($L_{Grad-CAM}^c$) é obtido pela combinação linear ponderada dos mapas de características A^k pelos pesos α_k^c, seguida pela aplicação da função ReLU para manter apenas as contribuições positivas (características que aumentam a pontuação da classe):
$$L_{Grad-CAM}^c = \text{ReLU}\left(\sum_k \alpha_k^c A^k\right)$$

O resultado é um mapa de calor que sobrepõe a imagem de entrada, destacando as regiões relevantes que a rede utilizou para tomar sua decisão de classificação para a classe escolhida. Essa visualização fornece uma forma intuitiva de validar a decisão do modelo, sendo útil em tarefas de classificação onde é importante saber se o modelo está considerando a região correta em vez de um viés no plano de fundo.

2.3 PLACENTAS

A placenta é um órgão transitório que se desenvolve exclusivamente durante a gestação e desempenha papel central na manutenção da gravidez. Ela atua como uma interface entre os sistemas circulatórios da mãe e do feto, sendo responsável por funções vitais como o transporte de nutrientes e oxigênio, remoção de resíduos metabólicos, produção hormonal e modulação imunológica.

Alterações estruturais ou funcionais na placenta podem comprometer o desenvolvimento fetal e estão associadas a diversas complicações obstétricas, como pré-eclâmpsia, restrição de crescimento intrauterino, parto prematuro e óbito fetal. Dessa forma, a avaliação morfológica placentária fornece informações diagnósticas e prognósticas valiosas tanto para o desfecho da gestação atual quanto para gestações futuras.

2.3.1 Anatomia placentária

A placenta humana é do tipo hemo-corial, ou seja, o sangue materno entra em contato direto com o tecido fetal. Sua estrutura pode ser dividida em duas porções principais:

- **Porção materna:** voltada para a parede uterina, apresenta aspecto rugoso e lobulado. É composta por múltiplos lóbulos chamados cotilédones, que correspondem a subdivisões do espaço intervilositário, onde ocorre o banho do sangue materno sobre as vilosidades fetais.
- **Porção fetal:** voltada para o feto, é recoberta pelas membranas fetais (amniótica e coriônica) e conecta-se ao feto por meio do cordão umbilical. Apresenta uma superfície lisa, brilhante e com vasos sanguíneos visíveis irradiando a partir da inserção do cordão. Essa face é formada principalmente por vilosidades coriônicas, estruturas ramificadas que aumentam a área de contato para as trocas materno-fetais.

Figura 11 – Imagens de uma placenta expondo a porção materna (esquerda) e fetal (direita).

Além dessas porções principais, a placenta é constituída por diversas estruturas funcionais importantes:

- **Cordão umbilical:** conecta o feto à placenta e contém duas artérias e uma veia, responsáveis por transportar sangue desoxigenado do feto para a placenta e sangue oxigenado da placenta para o feto, respectivamente.
- **Membranas fetais:** as membranas amniótica (interna) e coriônica (externa) envolvem o feto e o líquido amniótico, protegendo e sustentando o ambiente intrauterino.
- **Espaço intervilositário:** local onde o sangue materno circula livremente ao redor das vilosidades fetais, permitindo as trocas gasosas e nutricionais entre os dois organismos.
- **Vilosidades coriônicas:** unidades estruturais funcionais da placenta, ricamente vascularizadas, por onde ocorrem as trocas metabólicas.

2.3.2 Alterações morfológicas

Alterações morfológicas na placenta referem-se a modificações visíveis em sua estrutura, forma, coloração ou textura, identificáveis durante a inspeção macroscópica ou por meio de exame anatomopatológico. Essas alterações podem refletir processos patológicos ocorridos durante a gestação ou estarem associadas a desfechos clínicos adversos, como pré-eclâmpsia, restrição de crescimento intrauterino, parto prematuro ou óbito fetal.

A avaliação morfológica da placenta deve realizada por meio da análise de ambas as faces - materna e fetal - além do cordão umbilical e das membranas ovulares. Os seguintes aspectos são comumente observados em exame clínico e patológico:

- **Inserção marginal do cordão umbilical com distribuição radial dos vasos:** ocorre quando o cordão se insere na borda da placenta, com vasos se distribuindo radialmente a partir da inserção. Pode estar associada a menor eficiência nas trocas materno-fetais.
- **Inserção periférica com vascularização radial:** corresponde à inserção próxima à borda do disco placentário, porém com menor grau de deslocamento em relação à margem quando comparada à inserção marginal. Pode estar relacionada a alterações na implantação placentária.
- **Disco placentário irregular ou oval:** o formato anormal do disco pode sugerir alterações no desenvolvimento placentário ou má adaptação uterina.
- **Membranas ovulares opacas:** podem indicar processos inflamatórios ou infecciosos que afetam as membranas fetais, como a corioamnionite.
- **Lesões placentárias:** incluem áreas com sinais de infarto, necrose, calcificação ou hemorragia. Representam importantes marcadores de sofrimento placentário e comprometimento fetal.

- **Membranas ovulares esverdeadas:** normalmente associadas à presença de mecônio - substância eliminada pelo intestino do recém-nascido - no líquido amniótico, o que pode indicar sofrimento fetal crônico ou agudo.
- **Placenta retida:** condição em que a placenta ou fragmentos placentários não são expelidos espontaneamente após o parto, podendo causar hemorragia e infecção.
- **Descolamento prematuro da placenta:** separação parcial ou total da placenta antes do nascimento do bebê, com risco elevado para mãe e feto.
- **Cordão umbilical hiperenovelado:** aumento do número de espirais por 10 cm de cordão (índice de torção aumentado), alteração que pode modificar a hemodinâmica fetoplacentária e, em alguns casos, associar-se a hipóxia fetal ou restrição do crescimento.
- **Nós verdadeiros no cordão umbilical:** formação de laços completos no cordão durante a gestação, podendo, se apertados, reduzir o suprimento sanguíneo fetal.

Figura 12 – Exemplo de uma placenta morfologicamente alterada, com um nó verdadeiro no cordão umbilical.

3 ESTADO DA ARTE

Neste capítulo apresenta-se o estado da arte dos modelos e técnicas de inteligência artificial aplicados para a classificação de imagens de placentas.

3.1 METODOLOGIA DE REVISÃO SISTEMÁTICA DE LITERATURA

A revisão sistemática de literatura foi realizada com base em procuras de palavras-chave, conforme o método descrito por Kitchenham (KITCHENHAM, 2004). O objetivo desta etapa é estabelecer quais são os métodos de visão computacional utilizados na identificação de alterações macroscópicas placentárias.

3.1.1 Bancos de dados de pesquisa

Foram considerados jornais, revistas e conferências do ano de 2019 a 2024. Foram utilizadas as seguintes fontes para a pesquisa:

- IEEE Xplore
- ScienceDirect
- ACM Digital Library
- Pubmed
- Scielo

3.1.2 Termos de busca

Foram utilizadas as seguintes palavras-chave para a busca dos artigos:

(placenta) AND (pathology OR placental pathology OR macroscopic OR gross pathology OR gross morphology) AND (deep learning OR Convolutional neural network OR CNN OR Transfer learning OR Machine Learning OR Segmentation) AND (Photo image analysis)

3.1.3 Procedimento de seleção

No processo de seleção dos estudos, foram selecionados aqueles que atendiam a critérios mínimos de qualidade e inclusão, sem atender a nenhum dos critérios de exclusão:

Critérios de qualidade:

- Apresentam descrição detalhada da metodologia, incluindo arquitetura dos modelos de *deep learning*, datasets utilizados e métricas de avaliação.

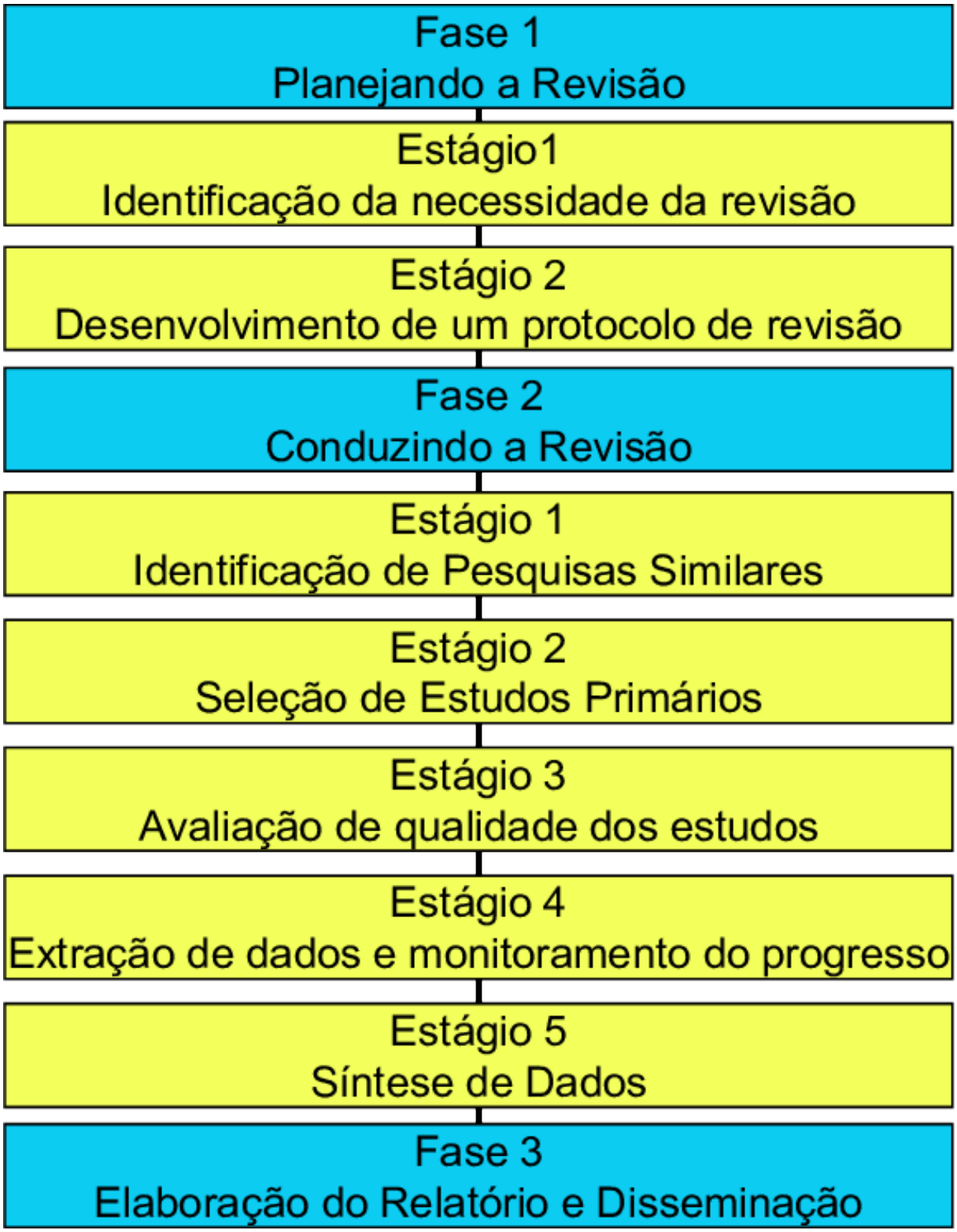


Figura 13 – Resumo da metodologia de revisão sistemática da literatura, descrito por Kitchenham (PEREIRA DA SILVA, 2009).

- Contêm validação experimental com dados reais, ou seja, testes realizados em imagens de placentas físicas, e não apenas em simulações ou modelos sintéticos.
- Utilizam técnicas de segmentação, detecção ou classificação de padrões morfológicos placentários com base em imagens fotográficas.
- Discutem limitações do método proposto, incluindo desafios enfrentados na detecção de padrões morfológicos.
- Realizam comparação quantitativa com outras abordagens do estado da arte, utilizando métricas padronizadas como IoU, Dice Score, AUC-ROC, mAP.
- Publicados em revistas ou conferências de alto impacto, indexadas em bases científicas reconhecidas.

Critérios de inclusão:

- Publicação em língua inglesa e disponível na íntegra nas bases de dados científicas selecionadas.
- Publicação no período 2019 a 2024, garantindo que os estudos reflitam os avanços recentes no uso de *deep learning* para análise placentária.
- Aplicação de redes neurais profundas para análise morfológica macroscópica de placentas por meio de imagens fotográficas do órgão.
- Estudos que incluem validação e experimentação prática, com uso de datasets reais e apresentação de métricas quantitativas para avaliação dos modelos.
- Estudos que consideram aspectos clínicos relevantes, como precisão no diagnóstico de alterações placentárias.

Critérios de exclusão:

- Estudos que analisam alterações macroscópicas placentárias exclusivamente por exames de imagem clínica, como ressonância magnética, ultrassonografia ou tomografia computadorizada, sem o uso de imagens fotográficas do órgão.
- Trabalhos que são apenas propostas teóricas de modelos, sem validação experimental ou resultados demonstrados.
- Artigos que não detalham métricas de avaliação ou não apresentam resultados concretos sobre o desempenho dos modelos utilizados.
- Estudos que não utilizam aprendizado profundo como abordagem principal para a análise placentária.

- Revisões sistemáticas ou artigos de revisão sem dados originais ou novos experimentos.
- Trabalhos publicados em eventos de baixa relevância científica, sem revisão por pares ou sem reconhecimento em bases indexadas.
- Trabalhos publicados antes de 2019.

3.2 TRABALHOS RELACIONADOS

3.2.1 PlacentaNet: Automatic Morphological Characterization of Placenta Photos with Deep Learning

O PlacentaNet (CHEN; WU *et al.*, 2019), é um sistema de análise automática de imagens fotográficas de placentas com foco em tarefas morfológicas essenciais: segmentação do disco placentário, classificação do lado fetal ou materno e localização do ponto de inserção do cordão umbilical. A base de dados usada no estudo incluiu 1.003 imagens de placentas, coletadas no *Northwestern Memorial Hospital*, com anotações detalhadas em nível de pixel para disco, cordão, régua e fundo.

A arquitetura do modelo é formada por uma estratégia multi-tarefa que compartilha um *encoder* comum que extrai pirâmides de características visuais (f1 a f5). Três decodificadores específicos são conectados: o *SegDecoder*, para segmentação por classes (placenta, cordão, régua, fundo); a *Classification Subnet*, para distinguir entre o lado fetal ou materno; e o *IPDecoder*, para localizar o ponto de inserção do cordão umbilical.

A formação do modelo seguiu uma abordagem de transferência de aprendizagem em duas etapas. Inicialmente, o *encoder* e o *SegDecoder* foram treinados do zero. Posteriormente, com o encoder fixado, a *Classification Subnet* e o *IPDecoder* foram treinados separadamente, incorporando técnicas de aumento de dados.

Como resultado, o PlacentaNet apresentou uma precisão de 97,51% na tarefa de classificação fetal/materno, superando 95,52% obtidos sem o *encoder* compartilhado. Já na localização do ponto de inserção, obtiveram acurácia por pixel de 98,73%, superando modelos como U-Net (98,10%) e SegNet (96,5%). Além das tarefas citadas, o artigo evidenciou os resultados para aplicações clínicas: para detecção de placenta retida, a rede obteve AUC de 0,836; e para classificação do tipo de inserção do cordão (central, excêntrico ou marginal), alcançou acurácia de 88%.

Em síntese, o PlacentaNet representa uma contribuição relevante ao aplicar aprendizado profundo multi-tarefa em imagens macro de placentas, melhorando a segmentação, classificação e localização em comparação a métodos isolados. O uso eficiente de *transfer learning* fortalece a aprendizagem de tarefas menos supervisionadas. Embora promissor para triagem automatizada pós-parto, o autor aponta futuro trabalho em expansão do *da-*

taset, validação multi-institucional e integração com diagnóstico patológico para robustecer o potencial clínico do sistema.

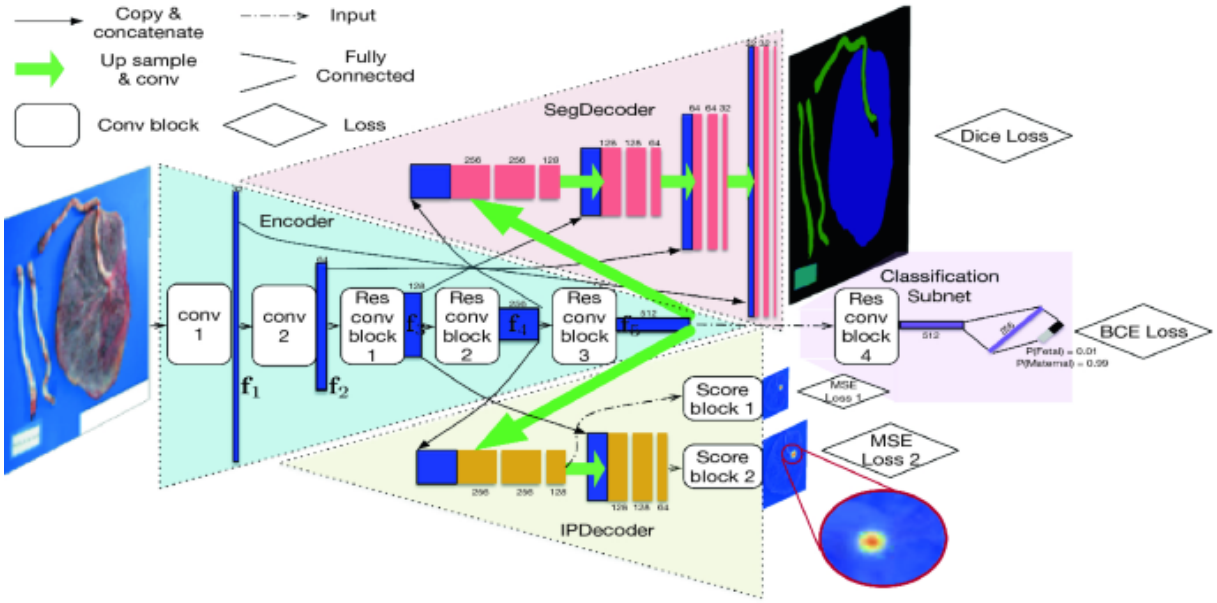


Figura 14 – Arquitetura do PlacentaNet (CHEN; WU *et al.*, 2019).

3.2.2 AI-PLAX: AI-based placental assessment and examination using photos

Em estudo subsequente (CHEN; ZHANG *et al.*, 2020), foi realizada a expansão do modelo proposto na seção 3.2.1. Nesse estudo, os autores coletaram um dataset de 1,370 imagens do *Northwestern Memorial Hospital*, capturadas entre os anos de 2016 e 2018.

A primeira etapa consiste de três redes *encoder-decoder* compartilhadas para a realização das tarefas de segmentação das estruturas da placenta, classificação do lado materno e fetal, e localização do ponto de inserção do cordão umbilical. O método proposto foi o detalhado na seção 3.2.1.

A segunda etapa consiste de sub-redes especializadas que exploram os *outputs* do primeiro estágio, juntamente com a imagem original, para extrair caraterísticas clínicas relevantes:

- Detecção de placenta retida: treinamento de uma rede neural convolucional ResNet-18 para a classificação binária de placenta retida com base nos mapas de segmentação do primeiro estágio. Obtenção de um resultado de AUC de 0,836, comparado a 0,781 da UNet e 0,844 da SegNet. Treinamento de uma rede neural convolucional ResNet-101 para a localização da porção retida. Obtenção de um resultado de IOU de 0,636.

- Categorização do tipo de inserção do cordão umbilical (central, excêntrica ou marginal): realizada com base na segmentação da placenta, localização do ponto de inserção e extração da escala a partir de uma régua presente na imagem. A técnica

inclui etapas como recuperação da área do disco placentário, retificação e análise da régua para converter pixels em centímetros, medição dos eixos longo e curto da placenta simulando a prática patológica, e cálculo da distância entre o ponto de inserção e a borda mais próxima. A classificação é feita com base na razão entre essa distância e a média dos eixos, utilizando limiares otimizados no conjunto de treinamento. O modelo alcançou 88% de acurácia na classificação dos tipos de inserção no conjunto de teste

- Detecção de mecônio, descolamento e corioamnionite: os autores utilizaram uma abordagem baseada em três classificadores binários independentes, cada um implementado por uma rede neural convolucional de 6 camadas. Apenas a região do disco placentário segmentada foi usada como entrada para os modelos. A detecção de mecônio atingiu sensibilidade de 0,98 e especificidade de 0,97, enquanto a de descolamento obteve 0,72/0,72 e a de corioamnionite 0,70/0,69, respectivamente.

- Detecção de formato irregular: os autores desenvolveram uma métrica de irregularidade baseada na comparação entre o contorno do disco placentário segmentado e uma elipse ajustada. A segmentação da placenta é usada para gerar um mapa binário do disco, do qual são extraídos momentos estatísticos de primeira e segunda ordem que permitem calcular os eixos, centro e orientação da elipse ajustada. A seguir, são contabilizados os pixels da elipse, os pixels do disco fora da elipse e os pixels fora do disco, mas dentro da elipse, para a composição da métrica de irregularidade. O modelo obteve sensibilidade de 0,87 e especificidade de 0,97.

- Detecção de cordão umbilical hiper-enrolado: foi proposto um método automatizado baseado na contagem de sulcos (dobras) detectados na região do cordão. Utilizando detecção de bordas de Canny (CANNY, 1986) sobre a segmentação do cordão, foram extraídos segmentos de bordas que representam as dobras causadas pelo enrolamento e aplicadas regras geométricas para evitar a contagem duplicada de sulcos ou de sulcos não representativos. Com base nesse critério, o modelo atingiu sensibilidade de 0,85 e especificidade de 0,93.

- Detecção de nós verdadeiros no cordão umbilical: os autores utilizaram uma arquitetura YOLO (REDMON *et al.*, 2016) treinada com imagens rotuladas manualmente com caixas delimitadoras dos nós. Como entrada para o modelo, foi utilizada a imagem RGB concatenada com uma máscara binária da região do cordão gerada pela segmentação. Atingiu-se uma AUC-ROC de 0,93.

O estudo demonstra que as representações aprendidas na primeira etapa - especialmente as provenientes da tarefa de segmentação - beneficiam diretamente tarefas subsequentes, como a classificação entre os lados fetal e materno e a localização do ponto

de inserção do cordão umbilical. Na segunda etapa, ao utilizar tanto os resultados da primeira quanto as imagens originais da placenta, modelos independentes foram empregados com sucesso em diversas análises clínicas relevantes. Para trabalhos futuros, os autores destacam a intenção de explorar técnicas de aprendizado semi-supervisionado, aproveitando grandes volumes de imagens não rotuladas, além de expandir a análise automatizada para indicadores patológicos adicionais e mais refinados, aprofundando o potencial de um *pipeline* multiestágio.

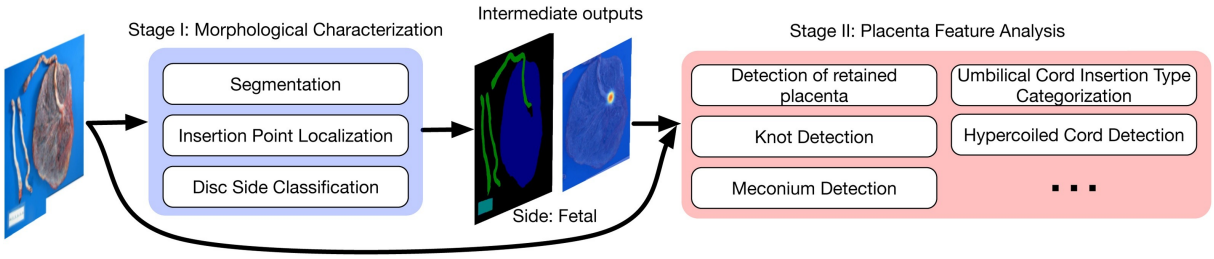


Figura 15 – Arquitetura expandida (CHEN; ZHANG *et al.*, 2020).

3.2.3 Multi-region saliency-aware learning for cross-domain placenta image segmentation

Em um trabalho subsequente (ZHANG *et al.*, 2020), foi desenvolvido um método de aprendizagem profunda denominado *Multi-Region Saliency-Aware Learning* (MSL), desenvolvido para melhorar a segmentação de imagens em diferentes domínios. O MSL emprega uma abordagem de adaptação de domínio ao integrar redes de atenção, geradores, discriminadores e restrições de consistência para a tradução de imagens placentárias entre domínios.

O resultado do MSL foi comparado com o de modelos estado da arte para a tradução de imagens, o CycleGAN e AGGAN. Ao adicionar adaptações apropriadas, verifica-se que a performance subiu de 0,2049 para 0,7380 no IoU médio. Além disso, o modelo obteve melhorias no diagnóstico de corioamnionite, exibindo uma melhoria de 0,78 e 0,76 para 0,96 e 0,83 no AUC de respostas inflamatórias fetais e maternas, respectivamente.

3.2.4 Vision-Language Contrastive Learning Approach to Robust Automatic Placenta Analysis Using Photographic Images

No artigo (PAN *et al.*, 2022) foi proposta a utilização da abordagem de aprendizado contrastivo de visão e linguagem (VLC) para analisar imagens de placentas em conjunto com os relatórios patológicos associados. Com isso, os autores buscam superar problemas de escassez de dados e melhorar a robustez e generalização do modelo.

O método utilizou ResNet-50 para o *encoder* de imagens e um modelo BERT pré-treinado para o encoder de texto. As imagens foram pré-processadas utilizando a segmentação do PlacentaNet - seção 3.2.1 -, além de realizarem a aplicação de aumentação

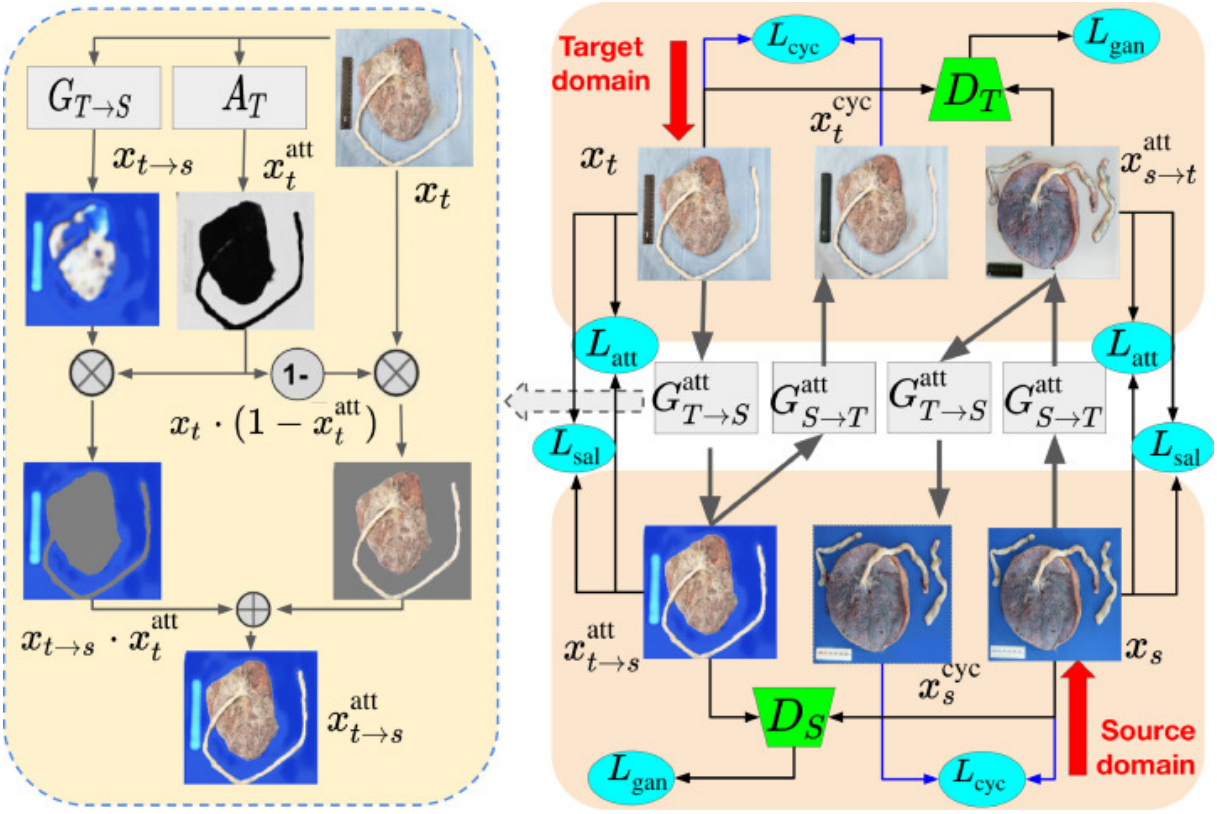


Figura 16 – *Pipeline* MSL proposto (ZHANG *et al.*, 2020).

de dados aleatórias. Por fim, as camadas de projeção de cada encoder foram conectadas com camadas totalmente conectadas para a realização da classificação.

Como base comparativa, os autores buscaram validar a arquitetura com uma ResNet-50 e com o *framework* ConVIRT para a classificação de alterações placentárias relacionadas ao mecônio, a respostas inflamatórias fetais, a respostas inflamatórias maternas, à corioamnionite e à sepse. A comparação demonstra aumentos de 1,4 a 11,3 pontos percentuais em relação às outras arquiteturas levantadas para imagens profissionais e para fotografias obtidas de *iPads*.

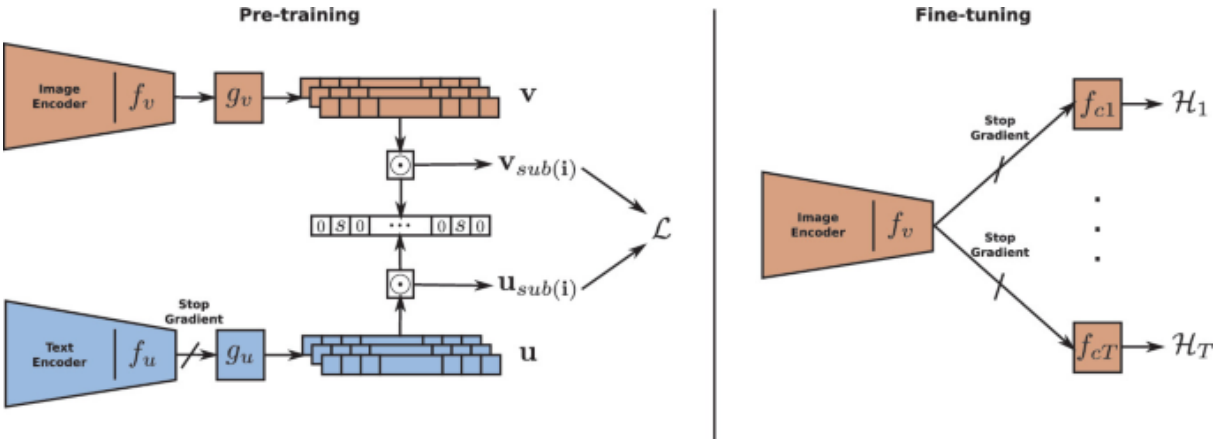


Figura 17 – Diagrama representativo do método VLC proposto (PAN *et al.*, 2022).

3.2.5 Enhancing Automatic Placenta Analysis through Distributional Feature Recomposition in Vision-Language Contrastive Learning

O artigo (PAN *et al.*, 2023) introduz algumas melhorias ao modelo VLC anterior, ditado na seção 3.2.4. A primeira melhoria é a recomposição de características do relatório de patologia, que busca resolver os problemas de supressão de características - já que o *encoder* pode ignorar certas *features* importantes das placentas para a diminuição da perda - e de comprimento do relatório, que pode ultrapassar a capacidade do *encoder*. A segunda melhoria foi a recomposição de características distribuicionais, que busca corrigir problemas de representação de features no espaço vetorial.

Para a implementação, foi utilizado um BERT pré-treinado como *encoder* de texto. Para as imagens, foi utilizado o PlacentaNet para o pré-processamento de imagens, com aumentação de dados aleatória e implementações padrão da EfficientNet-B0 e MobileNetV3-Large-1.0.

Como forma de verificação de resultados foi realizada uma comparação com três *baselines*: uma rede ResNet-50 de classificação, o *framework* médico VLC ConVIRT e a implementação anterior de Pan et. al. Foram verificadas melhorias entre 1,5 e 11,2 pontos percentuais em relação às baselines para a classificação de placentas com mecônio, corioamnionite histológica e clínica.

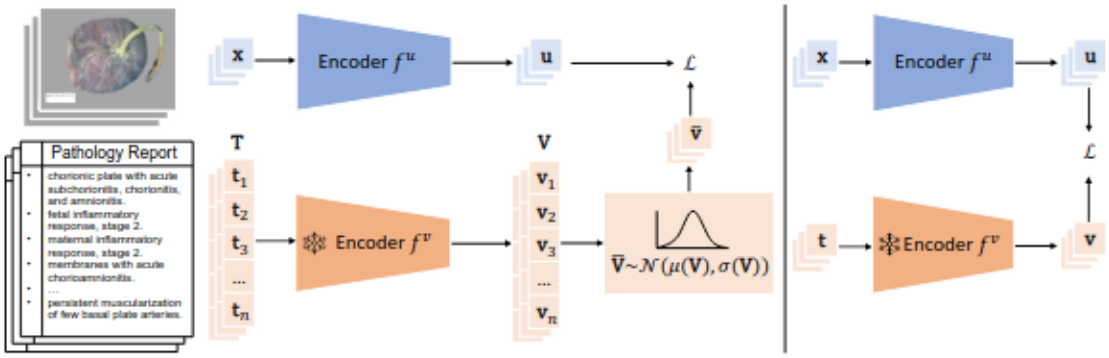


Figura 18 – Diagrama ilustrando a diferença entre o método proposto (esquerda) e o método VLC tradicional (direita) (PAN *et al.*, 2023).

4 DESENVOLVIMENTO

Este capítulo apresenta a implementação prática do *software* utilizado para o treinamento e a validação de redes neurais convolucionais na classificação de placentas com e sem alterações morfológicas (seção 4.2) e na classificação por tipo de alteração (seção 4.3).

Cada seção está dividida nas subseções:

- Conjunto de dados: descrição do *dataset* utilizado no desenvolvimento.
- Redes neurais: detalhamento das arquiteturas empregadas.
- Hiperparâmetros: apresentação dos hiperparâmetros de treinamento configurados.
- Ambiente computacional: especificação do ambiente computacional utilizado para o desenvolvimento e para os testes.

Além disso, como o mesmo conjunto de imagens foi utilizado em ambas as classificações, foi criada uma seção específica para descrevê-lo, incluindo suas características e distribuição, conforme apresentado na seção 4.1.

4.1 CONJUNTO DE IMAGENS

Para a realização das tarefas de classificação de alterações morfológicas em placentas, foram registradas 185 fotografias obtidas em centros obstétricos e hospitais de maternidade. Parte dessas imagens foi coletada no serviço de Obstetrícia do Hospital Universitário Polydoro Ernani de São Thiago da Universidade Federal de Santa Catarina (HU-UFSC/EBSERH), seguindo um protocolo de fotografia padronizado (LOPES; SILVA; VON WANGENHEIM, 2025), enquanto outras foram adquiridas em diferentes instituições sem padronização.

Uma imagem padronizada é aquela capturada sob um protocolo rigoroso que visa minimizar a variação não-clínica nos dados. Isso inclui o uso de um fundo característico (como uma placa azul de acrílico limpa e de alto contraste), iluminação controlada, e um ângulo e enquadramento adequados, com a placenta limpa e disposta ao centro da superfície. Em contraste, uma imagem não padronizada não segue nenhum rigor quanto à cor ou material do fundo, à limpeza da placenta, à sua disposição na superfície, ou às condições de iluminação.

As 185 imagens aptas para o estudo foram distribuídas conforme apresentado na tabela 1. Como se observa na tabela, o conjunto de dados apresenta um acentuado desequilíbrio de classes: apenas 31 imagens (16.8%) representam placentas com alterações morfológicas, enquanto 154 (83.2%) correspondem a placentas sem alterações. Esse desequilíbrio constitui um desafio metodológico relevante para o treinamento de modelos de

aprendizado profundo, exigindo estratégias específicas de mitigação, que serão abordadas nas seções seguintes. A figura 19 apresenta exemplos representativos de cada classe de placenta. As subfiguras a e h apresentam imagens que seguiram o protocolo de fotografia padronizado conforme descrito anteriormente, ao passo que as subfiguras b, c, d, e, f e g apresentam fotografias que não seguiram o protocolo.

Tabela 1 – Divisão das imagens do dataset por tipo de alteração.

Tipo	**Quantidade de imagens**	**Proporção**
Ausência de alteração	154	83.2%
Lesões placentárias	16	8.6%
Membranas ovulares opacas	4	2.2%
Inserção marginal do cordão umbilical com distribuição radial dos vasos sanguíneos	3	1.6%
Inserção periférica do cordão umbilical com vascularidade radial	3	1.6%
Disco placentário oval	2	1.1%
Membranas ovulares esverdeadas	2	1.1%
Disco placentário irregular	1	0.5%

Fonte: Autoria própria.

4.2 CLASSIFICAÇÃO DE PRESENÇA OU AUSÊNCIA DE ALTERAÇÕES MORFOLÓGICAS

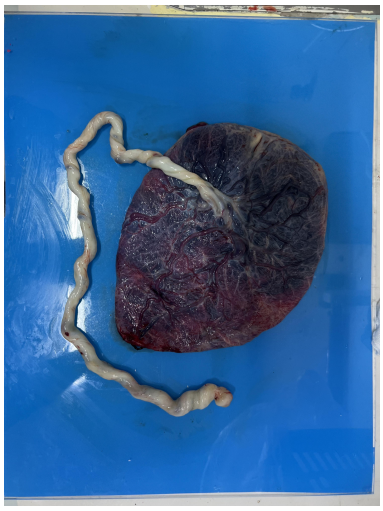
4.2.1 Conjunto de dados

Para essa tarefa, utilizou-se o conjunto de imagens descrito na seção 4.1, dividido em dois conjuntos: um submetido a técnicas de *data augmentation* (denominado “DA”) e outro sem essas transformações (denominado “noDA”).

Com o objetivo de mitigar o *overfitting* e ampliar a variabilidade dos dados, as transformações de *data augmentation* foram aplicadas exclusivamente ao conjunto de treino. Essa técnica não foi utilizada nos conjuntos de validação e teste, garantindo uma avaliação imparcial da capacidade de generalização dos modelos. As transformações aplicadas foram:

- Inversão (*flip*) horizontal: inversão da imagem horizontalmente com probabilidade de 50%.
- Rotação: rotação randômica da imagem entre -10° e +10°, aplicado com uma probabilidade de 75%.
- Ajuste de claridade e contraste: variação de até 10% em cada parâmetro, com uma probabilidade de 75%.

Figura 19 – Imagens de exemplo de cada classe de placentas.



(a) Ausência de alteração

(b) Disco placentário irregular

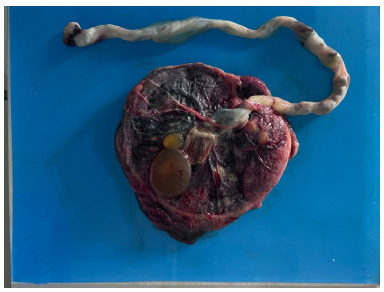
(c) Membranas ovulares opacas

(d) Inserção marginal do cordão umbilical com distribuição radial dos vasos

(e) Inserção periférica do cordão umbilical com vascularidade radial

(f) Disco placentário oval

(g) Membranas ovulares esverdeadas



(h) Lesões placentárias

- Cisalhamento (transformação de inclinação - *shear*): deformação de até 20% na imagem, com probabilidade de 75%.

O conjunto foi dividido em treino, validação e teste nas proporções de 80%, 10% e 10%, respectivamente. A tabela 2 apresenta a distribuição das imagens por classe em cada subconjunto.

Tabela 2 – Distribuição das imagens de placentas alteradas e não alteradas entre os conjuntos de treino, validação e teste

Classe	**Total**	**Treino**	**Validação**	**Teste**
Não alteradas	154	124	15	15
Alteradas	31	25	3	3
Total	**185**	**149**	**18**	**18**

Fonte: Autoria própria.

Por fim, as imagens foram redimensionadas para 224 × 224 *pixels*, padronizando o tamanho das entradas fornecidas às redes neurais.

4.2.2 Redes neurais

Foram avaliadas quatro arquiteturas de redes neurais convolucionais, utilizando a abordagem de aprendizagem por transferência (*transfer learning*). Todos os modelos foram inicializados com pesos pré-treinados no conjunto de dados ImageNet, o que permitiu aproveitar o aprendizado prévio de características visuais e acelerar o processo de convergência.

A seleção das arquiteturas buscou equilibrar modelos clássicos e robustos (ResNet) com arquiteturas modernas e eficientes (EfficientNet), de forma a comparar desempenho e complexidade. As arquiteturas avaliadas foram:

- ResNet-18
- ResNet-34
- EfficientNet-B0
- EfficientNet-B4

4.2.3 Hiperparâmetros

Os principais hiperparâmetros utilizados no treinamento estão descritos a seguir.

4.2.3.1 Função de perda

A função de perda adotada foi a *Cross Entropy Loss*, amplamente utilizada em tarefas de classificação, por medir a diferença entre a probabilidade prevista e a classe real, incentivando a maximização da probabilidade correta.

4.2.3.2 Otimizador

Utilizou-se o otimizador *Adam* (*Adaptive Moment Estimation*) (KINGMA; BA, 2015), que combina os benefícios do *Momentum* - acelerando a convergência - e do RMS-Prop - ajustando a taxa de aprendizado individualmente para cada parâmetro.

4.2.3.3 Política de um ciclo

A estratégia de variação cíclica da taxa de aprendizado e do *momentum*, conhecida como *One Cycle Policy* (SMITH; TOPIN, 2017), foi empregada para acelerar o treinamento e melhorar a generalização. A taxa máxima foi definida entre 1e-2 e 1e-3, com base no método *Learning Rate Finder* (*LR Finder*).

4.2.3.4 Tamanho do batch

O tamanho do lote foi definido por meio de uma busca em grade (*grid search*), avaliando valores de 8, 16 e 32 imagens.

4.2.3.5 Número de Épocas e parada antecipada

O treinamento foi configurado para até 100 épocas, com política de parada antecipada (*early stopping*) após cinco épocas consecutivas sem melhoria na validação. O modelo final corresponde ao de melhor desempenho no conjunto de validação.

4.2.4 Ambiente de desenvolvimento

Os experimentos foram conduzidos na plataforma *Google Colaboratory* (*Colab*), com suporte de uma unidade de processamento gráfico (GPU) NVIDIA Tesla T4.

O ambiente de *software* utilizou a linguagem Python e a biblioteca *fastai* para definição, treinamento e avaliação dos modelos de *deep learning*.

4.3 CLASSIFICAÇÃO POR TIPO DE ALTERAÇÃO MORFOLÓGICA

4.3.1 Conjunto de dados

Para a classificação do tipo de alteração morfológica, utilizou-se o conjunto de imagens descrito na seção 4.1. Como esse conjunto é pequeno e apresenta desequilíbrio entre classes, ele não seria suficiente para uma divisão adequada em conjuntos de treino,

validação e teste. Por esse motivo, aplicou-se a técnica de *data augmentation* para gerar novas imagens a partir das fotografias originais.

Essa estratégia tem como objetivo aumentar artificialmente a diversidade do conjunto de treino, aplicando transformações geométricas e fotométricas em tempo real. O pipeline de transformações, implementado com a biblioteca *Albumentations*, foi definido da seguinte forma:

- Inversão horizontal: aplicada com probabilidade de 50%.
- Inversão vertical: aplicada com probabilidade de 50%.
- Rotação de 90°: rotação aleatória em incrementos de 90°, com probabilidade de 50%.
- Transformação afim: aplicada com probabilidade de 50%, combinando:
  - Translação (*Shift*): deslocamento de até 5% da imagem.
  - Rotação (*Rotate*): rotação aleatória entre -15° e +15°.
  - A escala (*Scale*) foi mantida em 1.0 (sem zoom).
- Brilho e contraste: ajuste aleatório com probabilidade de 40%.
- Matiz e saturação: alteração aleatória de matiz, saturação e valor (HSV), com probabilidade de 40%.
- Desfoque: aplicação de um desfoque Gaussiano com probabilidade de 30%.

Em seguida, o conjunto de dados foi dividido em treino, validação e teste, nas proporções de 80%, 10% e 10%, respectivamente. A tabela 3 detalha a distribuição das imagens por classe em cada conjunto.

Por fim, todas as imagens foram redimensionadas para 224×224 *pixels*, garantindo padronização do tamanho das fotografias utilizadas como entrada nas redes neurais.

4.3.2 Redes neurais

Foram avaliadas quatro arquiteturas de redes neurais convolucionais, utilizando a abordagem de aprendizagem por transferência (*transfer learning*). Todos os modelos foram inicializados com pesos pré-treinados no conjunto de dados ImageNet, o que permitiu aproveitar o aprendizado prévio de características visuais e acelerar o processo de convergência.

A seleção das arquiteturas buscou equilibrar modelos clássicos e robustos (ResNet) com arquiteturas modernas e eficientes (EfficientNet), de forma a comparar desempenho e complexidade. As arquiteturas avaliadas foram:

- ResNet-18

Tabela 3 – Distribuição das imagens por tipo de alteração entre os conjuntos de treino, validação e teste.

Classe (Tipo)	**Total**	**Treino**	**Validação**	**Teste**
Ausência de alteração	616	492	62	62
Lesões placentárias	64	50	7	7
Membranas ovulares opacas	16	12	2	2
Inserção marginal do cordão umbilical com distribuição radial dos vasos sanguíneos	12	8	2	2
Inserção periférica do cordão umbilical com vascularidade radial	12	8	2	2
Disco placentário oval	8	6	1	1
Membranas ovulares esverdeadas	8	6	1	1
Disco placentário irregular	4	2	1	1
Total	**740**	**584**	**78**	**78**

Fonte: Autoria própria.

- ResNet-34
- EfficientNet-B0
- EfficientNet-B4

4.3.3 Hiperparâmetros

Os principais hiperparâmetros utilizados no treinamento estão descritos a seguir.

4.3.3.1 Função de perda

A função de perda adotada foi a *Cross Entropy Loss*, amplamente utilizada em tarefas de classificação, por medir a diferença entre a probabilidade prevista e a classe real, incentivando a maximização da probabilidade correta.

4.3.3.2 Otimizador

Utilizou-se o otimizador *Adam* (*Adaptive Moment Estimation*) (KINGMA; BA, 2015), que combina os benefícios do *Momentum* - acelerando a convergência - e do RMS-Prop - ajustando a taxa de aprendizado individualmente para cada parâmetro.

4.3.3.3 Política de um ciclo

A estratégia de variação cíclica da taxa de aprendizado e do *momentum*, conhecida como *One Cycle Policy* (SMITH; TOPIN, 2017), foi empregada para acelerar o treinamento

e melhorar a generalização. A taxa máxima foi definida entre 1e-2 e 1e-3, com base no método *Learning Rate Finder* (*LR Finder*).

4.3.3.4 Tamanho do batch

Foi utilizado um tamanho de *batch* (lote) de 8 imagens.

4.3.3.5 Número de Épocas e parada antecipada

O treinamento foi configurado para até 100 épocas, com política de parada antecipada (*early stopping*) após cinco épocas consecutivas sem melhoria na validação. O modelo final corresponde ao de melhor desempenho no conjunto de validação.

4.3.4 Ambiente de desenvolvimento

Os experimentos foram conduzidos na plataforma *Google Colaboratory* (*Colab*), com suporte de uma unidade de processamento gráfico (GPU) NVIDIA Tesla T4.

O ambiente de *software* utilizou a linguagem Python e a biblioteca *fastai* para definição, treinamento e avaliação dos modelos de *deep learning*.

5 RESULTADOS

Este capítulo apresenta uma análise detalhada dos resultados obtidos com a implementação e validação das redes neurais para: (i) classificação da presença ou ausência de alterações morfológicas em placentas (5.1) e (ii) classificação do tipo de alteração morfológica em placentas (5.2). Por fim, a seção 5.3 discute os resultados obtidos em ambos os experimentos.

5.1 REDES NEURAIS DE CLASSIFICAÇÃO DE PRESENÇA OU AUSÊNCIA DE ALTERAÇÕES MORFOLÓGICAS EM PLACENTAS

As tabelas 4 e 5 apresentam os resultados das redes neurais ResNet-18, ResNet-34, EfficientNet-B0 e EfficientNet-B4 considerando batches de 8, 16 e 32. Cada tabela considera a acurácia, a precisão, o recall e o F1-Score obtidos para os conjuntos de validação e testes.

As figuras 20, 21 e 22 apresentam as matrizes de confusão correspondentes às diferentes arquiteturas e configurações, destacando os modelos que atingiram os melhores F1-Scores.

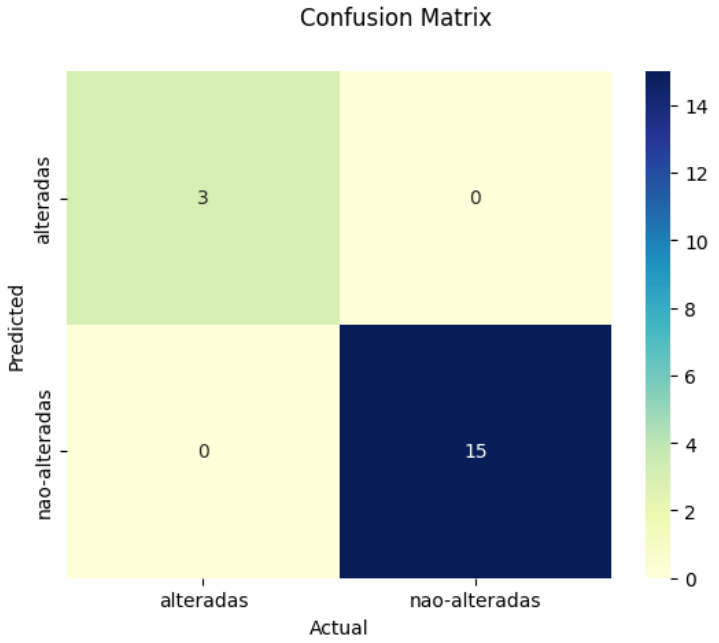


Figura 20 – Matriz de confusão para as configurações que atingiram um F1-Sore de **1.000**. Arquiteturas/configurações: ResNet-18 (NoDA, BS=8), EfficientNet-B4 (NoDA, BS=8), ResNet-18 (DA, BS=8), EfficientNet-B4 (DA, BS=8), ResNet-34 (NoDA, BS=16), EfficientNet-B4 (NoDA, BS=16), ResNet-18 (DA, BS=16), ResNet-34 (DA, BS=16), EfficientNet-B0 (DA, BS=16), EfficientNet-B4 (DA, BS=16), ResNet-18 (NoDA, BS=32), EfficientNet-B4 (DA, BS=32), ResNet-18 (DA, BS=32), ResNet-34 (DA, BS=32), EfficientNet-B4 (DA, BS=32).

Tabela 4 – Resultados experimentais na fase de validação, com e sem data augmentation.

Modelo	***Augmentation***	**Métrica**	***Batch* 8**	***Batch* 16**	***Batch* 32**
ResNet-18	NoDA	Acurácia	0.944	1.000	1.000
		Precisão	0.875	1.000	1.000
		Recall	0.967	1.000	1.000
		F1-Score	0.911	1.000	1.000
	DA	Acurácia	1.000	1.000	1.000
		Precisão	1.000	1.000	1.000
		Recall	1.000	1.000	1.000
		F1-Score	1.000	1.000	1.000
ResNet-34	NoDA	Acurácia	1.000	0.944	0.944
		Precisão	1.000	0.969	0.969
		Recall	1.000	0.833	0.833
		F1-Score	1.000	0.884	0.884
	DA	Acurácia	0.944	1.000	1.000
		Precisão	0.969	1.000	1.000
		Recall	0.833	1.000	1.000
		F1-Score	0.884	1.000	1.000
EfficientNet-B0	NoDA	Acurácia	1.000	1.000	1.000
		Precisão	1.000	1.000	1.000
		Recall	1.000	1.000	1.000
		F1-Score	1.000	1.000	1.000
	DA	Acurácia	1.000	1.000	1.000
		Precisão	1.000	1.000	1.000
		Recall	1.000	1.000	1.000
		F1-Score	1.000	1.000	1.000
EfficientNet-B4	NoDA	Acurácia	1.000	1.000	1.000
		Precisão	1.000	1.000	1.000
		Recall	1.000	1.000	1.000
		F1-Score	1.000	1.000	1.000
	DA	Acurácia	1.000	1.000	1.000
		Precisão	1.000	1.000	1.000
		Recall	1.000	1.000	1.000
		F1-Score	1.000	1.000	1.000

Tabela 5 – Resultados experimentais na fase de teste, com e sem data augmentation.

Modelo	***Augmentation***	**Métrica**	***Batch* 8**	***Batch* 16**	***Batch* 32**
ResNet-18	NoDA	Acurácia	1.000	0.944	0.889
		Precisão	1.000	0.875	0.800
		Recall	1.000	0.967	0.933
		F1-Score	1.000	0.911	0.839
	DA	Acurácia	1.000	1.000	1.000
		Precisão	1.000	1.000	1.000
		Recall	1.000	1.000	1.000
		F1-Score	1.000	1.000	1.000
ResNet-34	NoDA	Acurácia	0.944	1.000	1.000
		Precisão	0.875	1.000	1.000
		Recall	0.967	1.000	1.000
		F1-Score	0.911	1.000	1.000
	DA	Acurácia	0.944	1.000	0.944
		Precisão	0.875	1.000	0.875
		Recall	0.966	1.000	0.966
		F1-Score	0.911	1.000	0.911
EfficientNet-B0	NoDA	Acurácia	0.889	0.944	0.944
		Precisão	0.800	0.875	0.875
		Recall	0.933	0.967	0.967
		F1-Score	0.839	0.911	0.911
	DA	Acurácia	0.944	1.000	1.000
		Precisão	0.875	1.000	1.000
		Recall	0.967	1.000	1.000
		F1-Score	0.911	1.000	1.000
EfficientNet-B4	NoDA	Acurácia	1.000	1.000	1.000
		Precisão	1.000	1.000	1.000
		Recall	1.000	1.000	1.000
		F1-Score	1.000	1.000	1.000
	DA	Acurácia	1.000	1.000	1.000
		Precisão	1.000	1.000	1.000
		Recall	1.000	1.000	1.000
		F1-Score	1.000	1.000	1.000

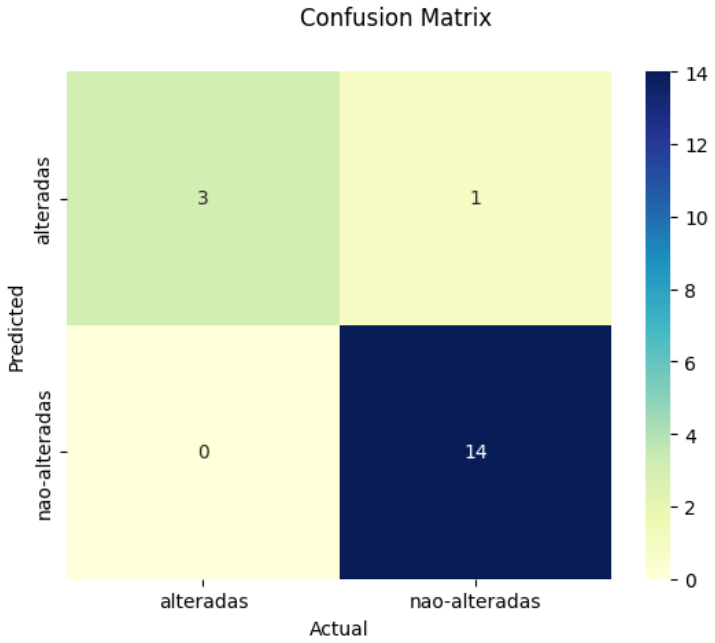


Figura 21 – Matriz de confusão para as configurações que atingiram um F1-Sore de **0.884**. Arquiteturas/configurações: ResNet-34 (NoDA, BS=8), ResNet-34 (DA, BS=8), EfficientNet-B0 (DA, BS=8), ResNet-18 (NoDA, BS=16), EfficientNet-B0 (NoDA, BS=16), EfficientNet-B0 (NoDA, BS=32), ResNet-34 (DA, BS=32).

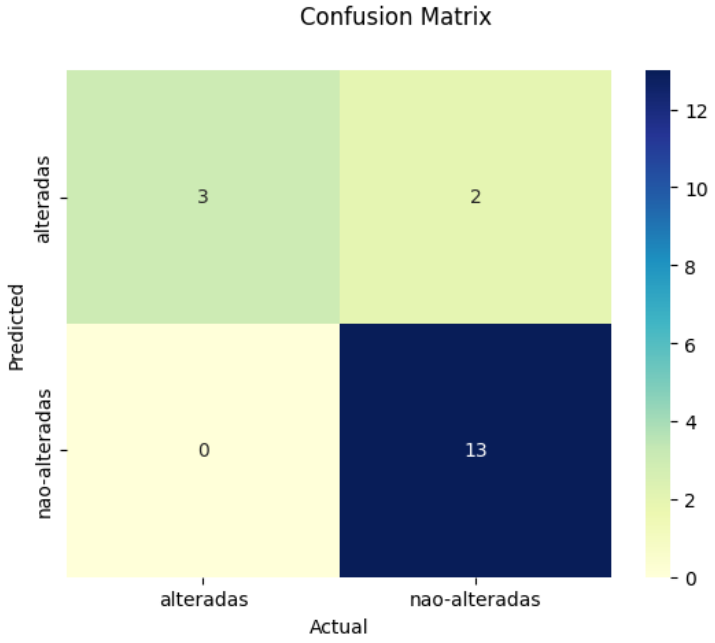


Figura 22 – Matriz de confusão para as configurações que atingiram um F1-Score de **0.839**. Arquiteturas/configurações: EfficientNet-B0 (NoDA, BS=32), ResNet-18 (NoDA, BS=32).

Os resultados indicam desempenho geral muito satisfatório para todas as arquiteturas, tanto na validação quanto no teste. Mesmo sem data augmentation (NoDA), as

métricas mostram que os modelos capturaram corretamente as classes.

Comparando as arquiteturas, EfficientNet apresentou resultados mais consistentes na validação. Na fase de teste, houve variação: ResNet-18 com data augmentation superou EfficientNet-B0 em alguns casos, indicando que não é possível afirmar que EfficientNet seja sempre superior nesta tarefa.

O uso de data augmentation beneficiou o treinamento, aumentando a variabilidade das imagens e melhorando os resultados, embora alguns casos isolados (como ResNet-34 com batch de 32) não tenham seguido esse padrão.

Quanto à profundidade, ResNet-18 teve desempenho superior à ResNet-34 com data augmentation, enquanto sem data augmentation a ResNet-34 foi melhor. Já na família EfficientNet, o modelo mais profundo (B4) apresentou desempenho perfeito em todas as métricas, mostrando que a profundidade favoreceu o resultado neste caso.

O tamanho do batch apresentou influência leve: batches maiores tendem a melhorar a estabilidade do treinamento, com batch de 16 apresentando desempenho mais positivo.

Em síntese, as arquiteturas EfficientNet-B4 (com ou sem data augmentation) e ResNet-18 com data augmentation mostraram o conjunto mais consistente de resultados.

5.2 REDES NEURAIS DE CLASSIFICAÇÃO DO TIPO DE ALTERAÇÃO MORFOLÓGICA EM PLACENTAS

As tabelas 6 e 7 apresentam os resultados de ResNet-18, ResNet-34, EfficientNet-B0 e EfficientNet-B4 com batch de 8, considerando acurácia, precisão, recall e F1-Score para validação e teste. As figuras 23, 24, 25 e 26 mostram as matrizes de confusão correspondentes.

Tabela 6 – Resultados experimentais na fase de validação.

Modelo	**Métrica**	**Resultado**
ResNet-18	Acurácia	1.000
	Precisão	1.000
	Recall	1.000
	F1-Score	1.000
ResNet-34	Acurácia	1.000
	Precisão	1.000
	Recall	1.000
	F1-Score	1.000
EfficientNet-B0	Acurácia	1.000
	Precisão	1.000
	Recall	1.000
	F1-Score	1.000
EfficientNet-B4	Acurácia	1.000
	Precisão	1.000
	Recall	1.000
	F1-Score	1.000

Tabela 7 – Resultados experimentais na fase de teste.

Modelo	**Métrica**	**Resultado**
ResNet-18	Acurácia	0.974
	Precisão	0.873
	Recall	0.873
	F1-Score	0.873
ResNet-34	Acurácia	0.949
	Precisão	0.742
	Recall	0.732
	F1-Score	0.736
EfficientNet-B0	Acurácia	0.974
	Precisão	0.935
	Recall	0.980
	F1-Score	0.947
EfficientNet-B4	Acurácia	0.974
	Precisão	0.996
	Recall	0.875
	F1-Score	0.915

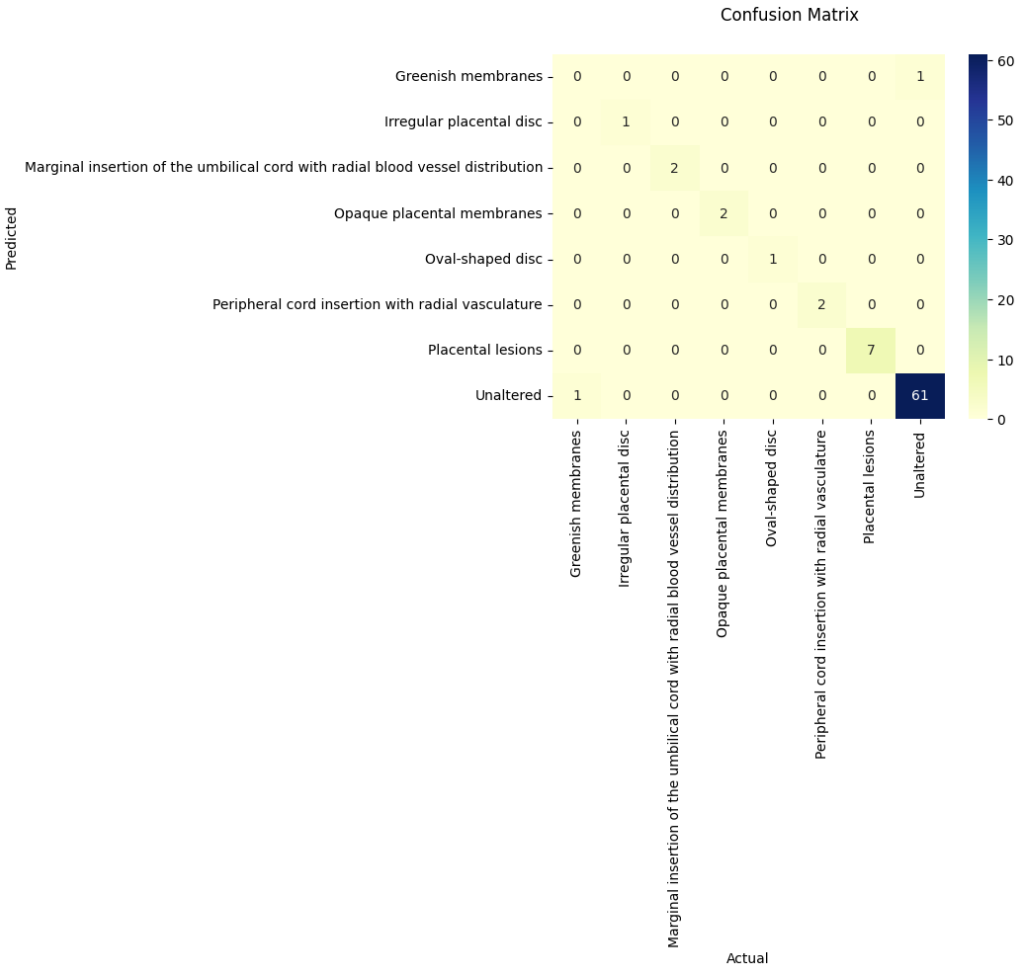


Figura 23 – Matriz de confusão para a arquitetura ResNet-18 na classificação do tipo de alteração em placentas.

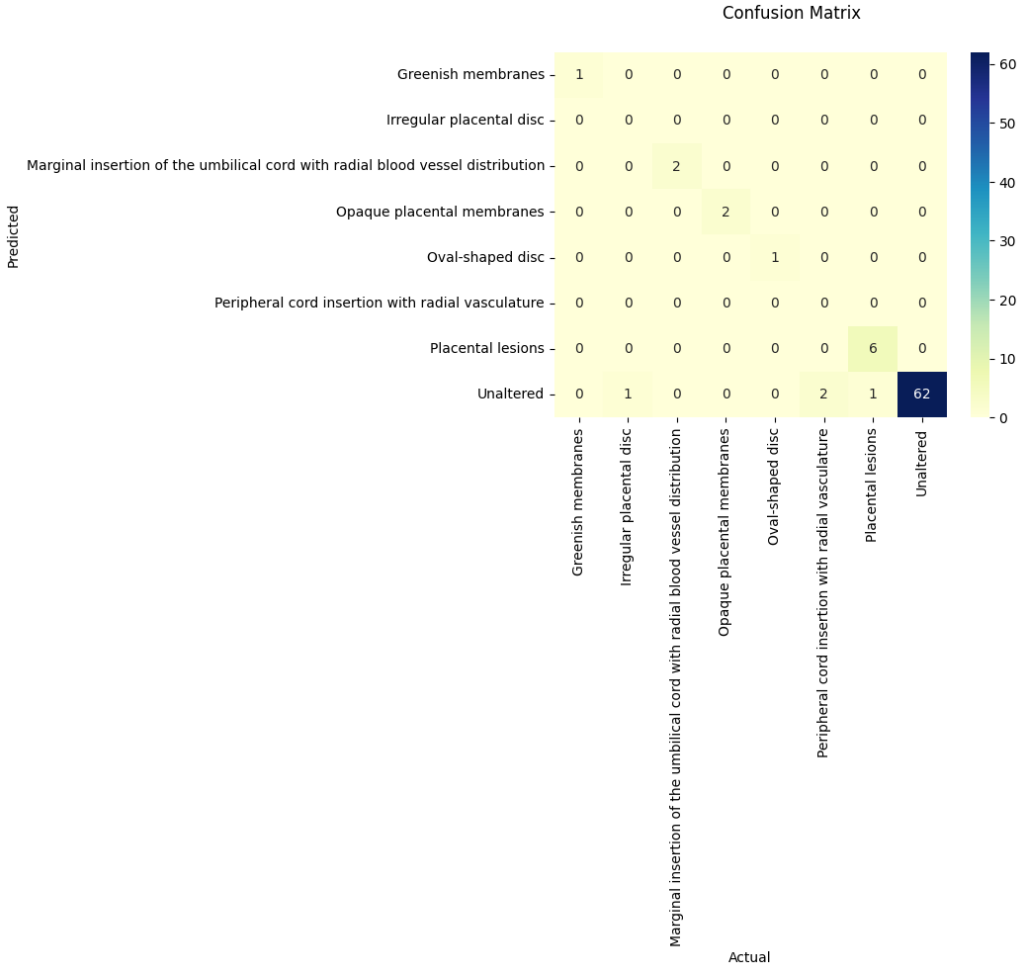


Figura 24 – Matriz de confusão para a arquitetura ResNet-34 na classificação do tipo de alteração em placentas.

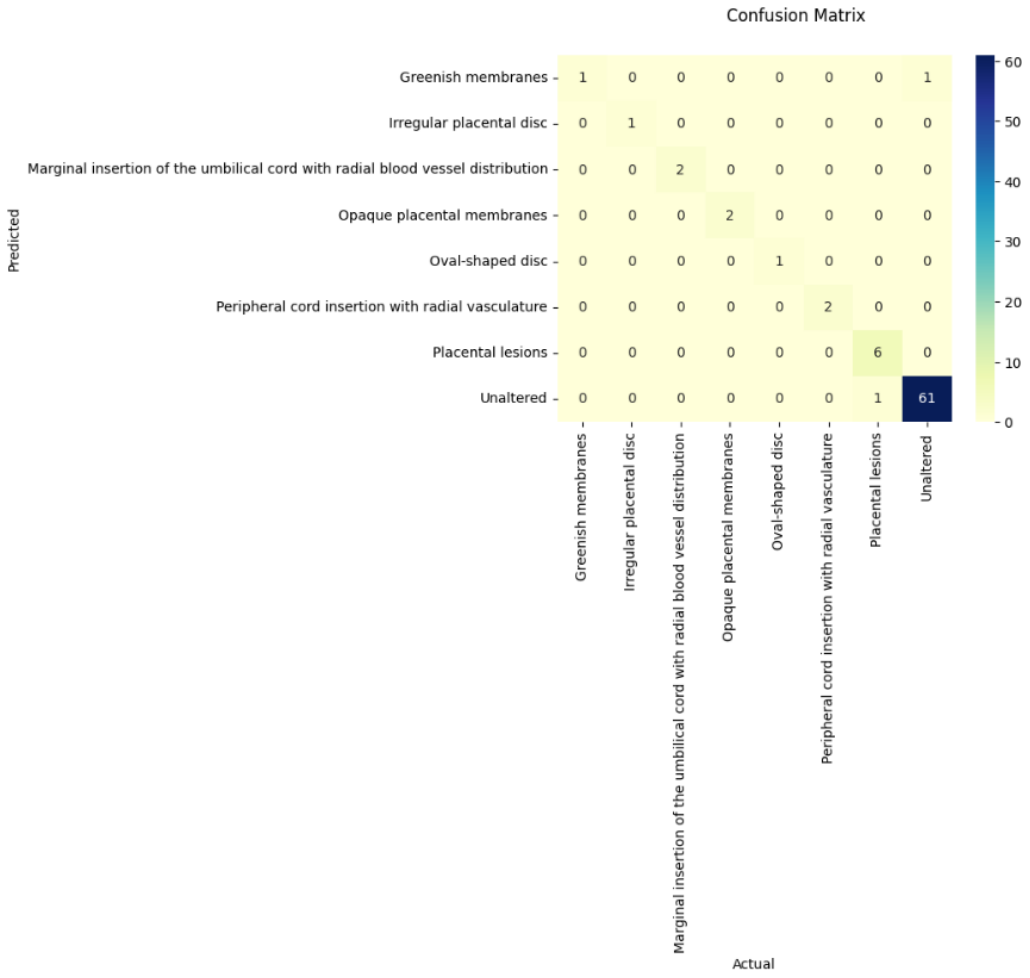


Figura 25 – Matriz de confusão para a arquitetura EfficientNet-B0 na classificação do tipo de alteração em placentas.

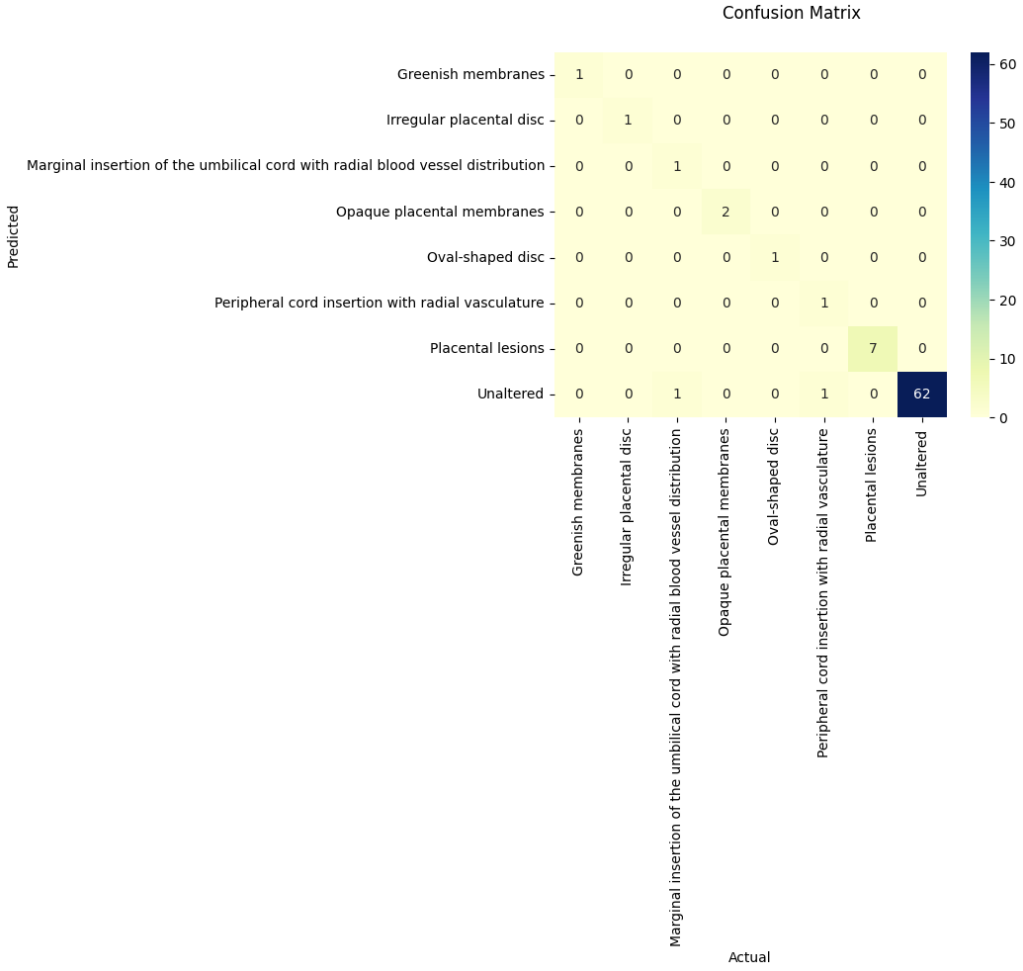


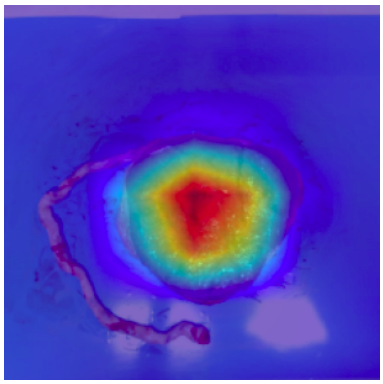
Figura 26 – Matriz de confusão para a arquitetura EfficientNet-B4 na classificação do tipo de alteração em placentas.

Observa-se que todas as redes obtiveram desempenho perfeito na validação. Na fase de teste, os resultados não foram máximos, mas permaneceram satisfatórios. O melhor desempenho no teste foi da EfficientNet-B0. Além disso, a EfficientNet apresentou métricas de precisão e recall superiores às de ResNet, provavelmente devido ao balanceamento otimizado entre profundidade, largura e resolução.

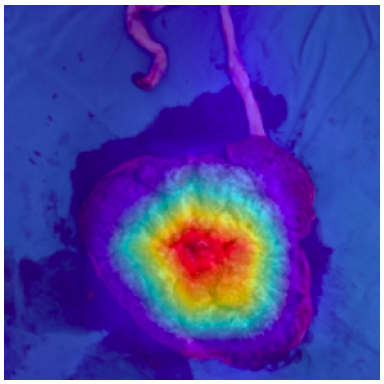
A figura 27 apresenta exemplos de visualizações obtidas pela técnica Grad-CAM, aplicada ao conjunto de teste da rede com melhor desempenho (EfficientNet-B0), a fim de ilustrar as regiões das placentas que foram consideradas mais relevantes para a classificação do tipo de alteração morfológica.

Pode-se verificar que principalmente a região central de cada placenta foi utilizada para a classificação dos tipos. Além disso, para a verificação das classes de inserção marginal ou periférica do cordão umbilical, pode-se observar que o foco da rede também foi no cordão umbilical.

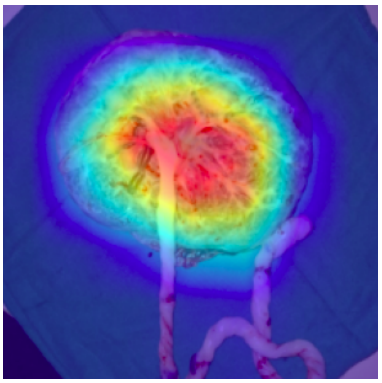
Figura 27 – Exemplos de visualizações Grad-CAM aplicadas a diferentes placentas no conjunto de teste, destacando as regiões de maior importância para a classificação do tipo de alteração morfológica.



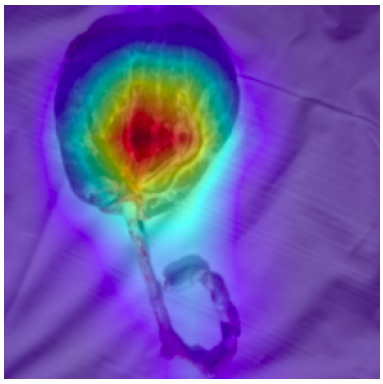
(a) Ausência de alteração



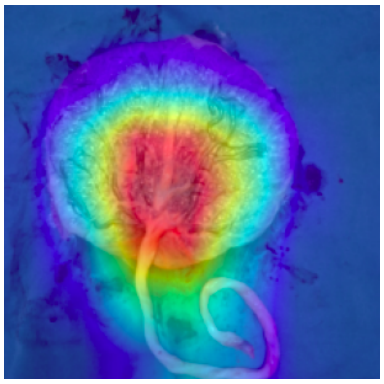
(b) Disco placentário irregular



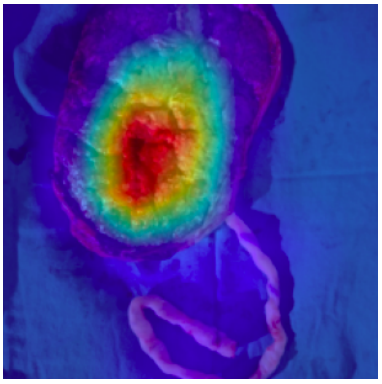
(c) Membranas ovulares opacas



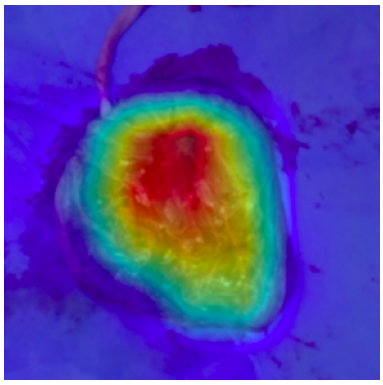
(d) Inserção marginal do cordão umbilical com distribuição radial dos vasos



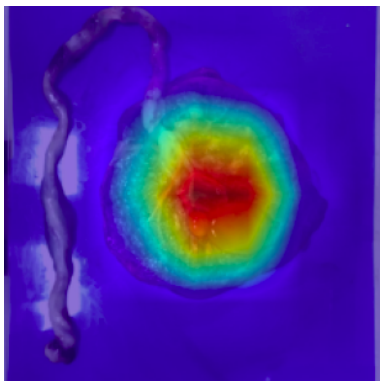
(e) Inserção periférica do cordão umbilical com vascularidade radial



(f) Disco placentário oval



(g) Membranas ovulares esverdeadas



(h) Lesões placentárias

5.3 DISCUSSÃO DOS RESULTADOS OBTIDOS

Os resultados demonstram sucesso em atingir os objetivos do estudo: classificação perfeita (100%) na presença ou ausência de alterações morfológicas e desempenho acima de 90% (97,4%) na classificação do tipo de alteração. Todas as métricas avaliadas confirmam a qualidade dos modelos treinados.

A principal limitação ao estudo realizado é o tamanho reduzido e o desbalanceamento do dataset. Muitas imagens apresentam variações mínimas de ângulo, iluminação ou perspectiva, aumentando o risco de *overfitting*, em que os modelos memorizam padrões específicos em vez de aprender representações generalistas. Assim, a performance observada pode superestimar a capacidade de predição em cenários clínicos variados.

6 CONCLUSÃO

Neste trabalho, foi empregado o método de trasfer learning para os modelos ResNet-18, ResNet-34, EfficientNet-B0 e EfficientNet-B4 com a finalidade de:

- Classificar a presença ou ausência de alterações morfológicas em placentas humanas
- Classificar o tipo de anomalia morfológica presente em placentas humanas

Os modelos foram avaliados com base em métricas relevantes, como acurácia, precisão, *recall* e F1-Score, demonstrando a eficácia das redes em ambas as tarefas.

Para a primeira tarefa, verificou-se que ResNet-18 com *data augmentation* e EfficientNet-B4 (com ou sem *data augmentation*) atingiram desempenho excelente, alcançando 100% em todas as métricas avaliadas. Na segunda tarefa, o modelo EfficientNet-B0 destacou-se, obtendo acurácia de 97,4% e F1-Score de 91,5%. Esses resultados evidenciam que o desempenho ótimo varia conforme a profundidade e o número de parâmetros das redes, bem como a natureza da tarefa.

O estudo também destacou a importância do *data augmentation* para contornar limitações do dataset, como tamanho reduzido e desbalanceamento de classes. Na primeira tarefa, a técnica contribuiu para resultados mais consistentes; na segunda, foi essencial para possibilitar uma divisão adequada do conjunto de dados e gerar variações suficientes das imagens.

Em síntese, este trabalho demonstrou a viabilidade do uso de modelos de aprendizado profundo na classificação de alterações morfológicas em placentas, atingindo com sucesso os objetivos gerais e específicos propostos.

6.1 TRABALHOS FUTUROS

Considerando as limitações constatadas neste trabalho, propõe-se a construção de um novo conjunto de dados contendo um número maior de amostras por classe e menor desbalanceamento entre elas, de modo a favorecer o aprendizado do modelo e aprimorar sua capacidade de generalização.

Além disso, recomenda-se realizar uma otimização mais aprofundada dos hiperparâmetros de treinamento, utilizando técnicas como *grid search* ou *Bayesian optimization* (SNOEK; LAROCHELLE; ADAMS, 2012), com o objetivo de identificar combinações que maximizem o desempenho das redes neurais.

Outro aspecto relevante a ser explorado é a investigação de arquiteturas mais recentes, como as baseadas em *Vision Transformers* (ViT) (DOSOVITSKIY *et al.*, 2020), que têm apresentado resultados promissores em tarefas de classificação de imagens médicas (CHEN; LU *et al.*, 2022).

Por fim, sugere-se a ampliação do escopo do estudo para incluir outras técnicas de explicabilidade (*explainable AI*) (TJOA; GUAN, 2021), permitindo compreender melhor as regiões de interesse utilizadas pelo modelo na tomada de decisão e, assim, contribuindo para maior transparência e confiabilidade em aplicações clínicas.

REFERÊNCIAS

ALANSARY, A. *et al.* Fast fully automatic segmentation of the human placenta from motion corrupted MRI. *In:* CONGRESSO Internacional de Processamento de Imagens Médicas e Intervenção Assistida por Computador. [*S.l.*: *s.n.*], 2016. P. 589–597.

ARAÚJO, C. P. *et al.* Alterações placentárias macroscópicas associadas à ocorrência de fetos/recebidos macrossômicos na gestação de alto risco. *In:* 58ª Reunião Anual da Sociedade Brasileira para o Progresso da Ciência. Florianópolis: SBPC, 2006. estado da arte.

CANNY, John. A Computational Approach to Edge Detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 8, n. 6, p. 679–698, 1986. DOI: 10.1109/TPAMI.1986.4767851.

CHANG, K. T. Pathological examination of the placenta: raison d'être, clinical relevance and medicolegal utility. **Singapore medical journal**, v. 50, p. 1123–1133, 2009.

CHEN, Jieneng; LU, Yongyi *et al.* TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. **arXiv preprint arXiv:2102.04306**, 2022.

CHEN, Yukun; WU, Chenyan *et al.* PlacentaNet: Automatic Morphological Characterization of Placenta Photos with Deep Learning. *In:* SHEN, Dinggang *et al.* (Ed.). **Medical Image Computing and Computer Assisted Intervention – MICCAI 2019**. Cham: Springer International Publishing, 2019. v. 11764. (Lecture Notes in Computer Science). P. 487–495. DOI: 10.1007/978-3-030-32239-7_54.

CHEN, Yukun; ZHANG, Zhuomin *et al.* AI-PLAX: AI-based placental assessment and examination using photos. **Computerized Medical Imaging and Graphics**, Elsevier, v. 84, p. 101744, set. 2020. Epub 2020 Jun 1. DOI: 10.1016/j.compmedimag.2020.101744. Disponível em: https://doi.org/10.1016/j.compmedimag.2020.101744.

DAS, D.; SADIQ, A. S.; MIRJALILI, S. Artificial Neural Networks: Structure and Learning. *In:* OPTIMIZATION Algorithms in Machine Learning. Singapore: Springer, 2025. P. 47–57.

DEL DUQUE, M. C. *et al.* O erro médico nos tribunais: uma análise das decisões do Tribunal de Justiça da capital brasileira. **Saúde Soc.**, v. 31, n. 3, 2022. DOI: 10.1590/S0104-12902022220144pt. Disponível em: https://doi.org/10.1590/S0104-12902022220144pt.

DOSOVITSKIY, Alexey *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. **arXiv preprint arXiv:2010.11929**, 2020. Disponível em: https://arxiv.org/abs/2010.11929.

GUIDOTTI, Riccardo *et al.* A survey of methods for explaining black box models. **ACM Computing Surveys (CSUR)**, ACM New York, NY, USA, v. 51, n. 5, p. 1–42, 2018.

HE, Kaiming *et al.* Deep Residual Learning for Image Recognition. *In:* PROCEEDINGS of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [*S.l.*: *s.n.*], jun. 2016. P. 770–778. DOI: 10.1109/CVPR.2016.90.

HOPE, Tom; RESHEFF, Yehezkel S.; LIEDER, Itay. **TensorFlow for Deep Learning: From Linear Regression to Reinforcement Learning**. [*S.l.*]: O'Reilly Media, 2018. ISBN 978-1491980453. Disponível em: https://www.oreilly.com/library/view/tensorflow-for-deep/9781491980446/.

HOWARD, Jeremy; GUGGER, Sylvain. fastai: A Layered API for Deep Learning. **Information**, v. 11, n. 2, p. 108, 2020. DOI: 10.3390/info11020108.

IBM. **What are Convolutional Neural Networks?** [*S.l.*: *s.n.*], 2025. Accessed: 2025-07-06. Disponível em: https://www.ibm.com/think/topics/convolutional-neural-networks.

KHONG, T. Y. *et al.* Sampling and Definitions of Placental Lesions: Amsterdam Placental Workshop Group Consensus Statement. **Arch Pathol Lab Med**, v. 140, n. 7, p. 698–713, 2016. DOI: 10.5858/arpa.2015-0225-CC.

KINGMA, Diederik P; BA, Jimmy. Adam: A Method for Stochastic Optimization. **International Conference on Learning Representations (ICLR)**, 2015. Disponível em: https://arxiv.org/abs/1412.6980.

KITCHENHAM, Barbara A. **Procedures for Performing Systematic Reviews**. Keele, UK, jul. 2004. Also NICTA report 0400011T.1. Disponível em: https://www.researchgate.net/publication/228756057_Procedures_for_Performing_Systematic_Reviews.

KUO, Ping-Huan; LIN, Ssu-Ting; HU, Jun. DNAE-GAN: Noise-free acoustic signal generator by integrating autoencoder and generative adversarial network. **International Journal of Distributed Sensor Networks**, v. 16, p. 155014772092352, mai. 2020. DOI: 10.1177/1550147720923529.

LANGSTON, C. *et al.* Practice guideline for examination of the placenta: developed by the Placental Pathology Practice Guideline Development Task Force of the College of American Pathologists. **Arch Pathol Lab Med**, v. 121, n. 5, p. 449–476, 1997.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015. DOI: 10.1038/nature14539.

LECUN, Yann; BOTTOU, Léon *et al.* Gradient-Based Learning Applied to Document Recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2324, nov. 1998. DOI:

10.1109/5.726791. Disponível em:
http://vision.stanford.edu/cs598_spring07/papers/Lecun98.pdf.

LITJENS, Geert *et al.* A survey on deep learning in medical image analysis. **Medical Image Analysis**, Elsevier, v. 42, p. 60–88, 2017.

LOONEY, P. *et al.* Automatic 3D ultrasound segmentation of the first trimester placenta using deep learning. *In:* IEEE International Symposium on Biomedical Imaging. [*S.l.*: *s.n.*], 2017. P. 279–282.

LOPES, Beatriz; SILVA, Bibiana; VON WANGENHEIM, Aldo. **PLACENTA IMAGE CAPTURE PROTOCOL IN THE DELIVERY ROOM/PROTOCOLO DE CAPTURA DE IMAGENS DE PLACENTAS NA SALA DE PARTO**. [*S.l.*], jan. 2025. DOI: 10.13140/RG.2.2.16008.94729.

LUNDBERG, Scott M; LEE, Su-In. A unified approach to interpreting model predictions. *In:* ADVANCES in neural information processing systems. [*S.l.*: *s.n.*], 2017. P. 4765–4774.

MALATHI, G.; SHANTHI, V. Statistical measurement of ultrasound placenta images complicated by gestational diabetes mellitus using segmentation approach. **J Inform Hiding Multimedia Signal Process**, p. 332–343, 2011.

MITTAL, N. *et al.* A practical guide to placental examination for forensic pathologists. **Forensic Science, Medicine and Pathology**, v. 16, p. 295–312, 2020. DOI: 10.1007/s12024-019-00214-2. Disponível em:
https://doi.org/10.1007/s12024-019-00214-2.

MUREL, Jacob; KAVLAKOGLU, Eda. **O que é aumento de dados?** [*S.l.*: *s.n.*], 2024. Acesso em: 6 jul. 2025. Disponível em:
https://www.ibm.com/br-pt/think/topics/data-augmentation.

PAN, Yiqiu *et al.* Enhancing Automatic Placenta Analysis Through Distributional Feature Recomposition in Vision-Language Contrastive Learning. *In:* GREENSPAN, Hayit *et al.* (Ed.). **Medical Image Computing and Computer Assisted Intervention - MICCAI 2023**. Cham: Springer, 2023. v. 14225. (Lecture Notes in Computer Science). P. 117–127. DOI: 10.1007/978-3-031-43987-2_12. Disponível em: https://doi.org/10.1007/978-3-031-43987-2_12.

_____. Vision-Language Contrastive Learning Approach to Robust Automatic Placenta Analysis Using Photographic Images. *In:* WANG, Liansheng *et al.* (Ed.). **Medical Image Computing and Computer Assisted Intervention - MICCAI 2022**. Cham: Springer, 2022. v. 13433. (Lecture Notes in Computer Science). P. 718–728. DOI: 10.1007/978-3-031-16437-8_68. Disponível em:
https://doi.org/10.1007/978-3-031-16437-8_68.

PEREIRA DA SILVA, Édison Renato. **Métodos para Revisão e Mapeamento Sistemático da Literatura (Methods for Systematic Literature Reviews and Systematic Mapping Studies)**. Mar. 2009. Tese (Doutorado). DOI: 10.13140/RG.2.1.2582.8723.

PTACEK, I. *et al.* Systematic review of placental pathology reported in association with stillbirth. **Placenta**, v. 35, n. 8, p. 552–562, 2014.

REDMON, Joseph *et al.* You Only Look Once: Unified, Real-Time Object Detection. *In:* PROCEEDINGS of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [*S.l.*: *s.n.*], 2016. P. 779–788. DOI: 10.1109/CVPR.2016.91.

RIBEIRO, Marco Tulio; SINGH, Sameer; GUESTRIN, Carlos. Why should I trust you?: Explaining the predictions of any classifier. *In:* PROCEEDINGS of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. [*S.l.*: *s.n.*], 2016. P. 1135–1144.

ROMERO, Miguel *et al.* Training Deep Learning models with small datasets, dez. 2019. DOI: 10.48550/arXiv.1912.06761.

SELVARAJU, Ramprasaath R *et al.* Grad-CAM: Visual explanations from deep networks via gradient-based localization. *In:* PROCEEDINGS of the IEEE international conference on computer vision. [*S.l.*: *s.n.*], 2017. P. 618–626.

SIMONYAN, Karen; VEDALDI, Andrea; ZISSERMAN, Andrew. Deep inside convolutional networks: Visualising image classification models and saliency maps. *In:* INTERNATIONAL conference on learning representations (ICLR) workshop. [*S.l.*: *s.n.*], 2014.

SMITH, Leslie N.; TOPIN, Nicholay. Super-Convergence: Very Fast Training of Neural Networks Using Large Learning Rates. **arXiv preprint arXiv:1708.07120**, 2017. Disponível em: https://arxiv.org/abs/1708.07120.

SNOEK, Jasper; LAROCHELLE, Hugo; ADAMS, Ryan P. Practical Bayesian optimization of machine learning algorithms. *In:* ADVANCES in Neural Information Processing Systems. [*S.l.*: *s.n.*], 2012. P. 2951–2959.

TAJBAKHSH, Nima *et al.* Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? **IEEE Transactions on Medical Imaging**, v. 35, n. 5, p. 1299–1312, 2016. DOI: 10.1109/TMI.2016.2535302.

TAN, Mingxing; LE, Quoc V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *In:* PROCEEDINGS of the 36th International Conference on Machine Learning. [*S.l.*]: PMLR, jun. 2019. P. 6105–6114.

TJOA, Erico; GUAN, Cuntai. A survey on explainable artificial intelligence (XAI): Toward medical XAI. **IEEE Transactions on Neural Networks and Learning Systems**, v. 32, n. 11, p. 4793–4813, 2021. DOI: 10.1109/TNNLS.2020.3027314.

WANG, Xuan; ZHU, Zhigang. Context understanding in computer vision: A survey. **Computer Vision and Image Understanding**, v. 229, p. 103646, 2023. ISSN 1077-3142. DOI: https://doi.org/10.1016/j.cviu.2023.103646. Disponível em: https://www.sciencedirect.com/science/article/pii/S1077314223000267.

YINGGE, Huo; ALI, Imran; LEE, Kang-Yoon. Deep Neural Networks on Chip - A Survey. *In:* p. 589–592. DOI: 10.1109/BigComp48618.2020.00016.

ZHANG, Zhuomin *et al.* Multi-region saliency-aware learning for cross-domain placenta image segmentation. **Pattern Recognition Letters**, v. 140, p. 165–171, dez. 2020. DOI: 10.1016/j.patrec.2020.10.004.

ZHOU, Bolei *et al.* Learning deep features for discriminative localization. *In:* PROCEEDINGS of the IEEE conference on computer vision and pattern recognition. [*S.l.*: *s.n.*], 2016. P. 2921–2929.

APÊNDICE A – REPOSITÓRIOS

O código desenvolvido para a realização do estudo pode ser encontrado nos repositórios abaixo:

- Repositório com código desenvolvido (Códigos@UFSC): `https://codigos.ufsc.br/stephan.krug/placental-classifier`
- Repositório com código desenvolvido (GitHub): `https://github.com/stephankruggg/placental-classifier`

APÊNDICE B – ARTIGO

Revista de Informática Teórica e Aplicada - RITA - ISSN 2175-2745
Vol. 32, Num. 3 (2025) 11-23

RESEARCH ARTICLE

Automatic Classification of Placental Anomalies Using Convolutional Neural Networks

Classificação Automática de Anomalias Placentárias Utilizando Redes Neurais Convolucionais

Stephan Krug[1]*, Beatriz Silva Lopes[2], Bibiana Quatrin Tiellet da Silva[3], Aldo von Wangenheim[4]

Abstract: This study proposes a computer vision approach to automatically classify the presence of macroscopic anomalies in human placentas. Four convolutional neural network (CNN) architectures — ResNet-18, ResNet-34, EfficientNet-B0, and EfficientNet-B4 — were evaluated under different hyperparameter configurations and with/without data augmentation techniques. The dataset comprised 185 placental images, including both standardized and non-standardized acquisitions, collected from deliveries performed at different Obstetric Centers and Maternity Hospitals. Among these, the images obtained at the University Hospital of the Federal University of Santa Catarina (HU-UFSC) followed a standardized acquisition protocol. The dataset was divided into two classes: altered and non-altered placentas. The models were trained using transfer learning and optimized with modern strategies such as the Adam optimizer and the One Cycle Policy.
The results showed outstanding performance, with F1-Scores reaching up to 100%, indicating the models' high ability to distinguish between normal and altered placentas. However, it is hypothesized that such high performance may be related to the dataset's limited diversity, potentially impacting the models' generalization capability. Therefore, future work will include new datasets to validate model robustness and enhance applicability in real-world clinical settings.
Keywords: Macroscopical placental analysis — Deep learning — Convolutional neural networks — Computer vision

Resumo: Este trabalho propõe uma abordagem baseada em visão computacional para classificar automaticamente a presença de anomalias macroscópicas em placentas humanas. Foram avaliadas quatro arquiteturas de redes neurais convolucionais — ResNet-18, ResNet-34, EfficientNet-B0 e EfficientNet-B4 — em diferentes cenários de hiperparâmetros e com/sem técnicas de data augmentation. O conjunto de dados foi composto por 185 imagens, incluindo aquisições padronizadas e não padronizadas, obtidas a partir de partos realizados em diferentes Centros Obstétricos e Maternidades. Entre elas, as imagens coletadas no Hospital Universitário da Universidade Federal de Santa Catarina (HU-UFSC) seguiram um protocolo de aquisição padronizado. O conjunto foi dividido em duas classes: placentas com alterações e sem alterações. As redes foram treinadas utilizando transfer learning e estratégias modernas de otimização, como o otimizador Adam e a política One Cycle.
Os resultados indicaram desempenho excepcional, com métricas de F1-Score chegando a 100%, sugerindo que os modelos são capazes de distinguir com alta precisão entre placentas normais e alteradas. No entanto, levanta-se a hipótese de que essa alta performance esteja relacionada à baixa diversidade do dataset, o que pode comprometer a capacidade de generalização dos modelos. Assim, estudos futuros incluirão novos conjuntos de dados para validar a robustez das redes treinadas e aprimorar sua aplicabilidade em cenários clínicos reais.
Palavras-Chave: Análise macroscópica de placentas — Aprendizado profundo — Redes neurais convolucionais — Visão computacional

[1] *Department of Computer Science and Statistics (INE), Federal University of Santa Catarina (UFSC), Brazil*
[2] *Department of Computer Science and Statistics (INE), Federal University of Santa Catarina (UFSC), Brazil*
[3] *Department of Computer Science and Statistics (INE), Federal University of Santa Catarina (UFSC), Brazil*
[4] *Department of Computer Science and Statistics (INE), Federal University of Santa Catarina (UFSC), Brazil*
*__Corresponding author__: stephan.krug@grad.ufsc.br
DOI: http://dx.doi.org/10.22456/2175-2745.XXXX • **Received:** dd/mm/yyyy • **Accepted:** dd/mm/yyyy
CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

The human placenta plays a vital role in fetal development by connecting the fetus to the maternal uterus. It facilitates the exchange of oxygen and nutrients, removes metabolic waste, and serves as an immunological barrier against infections and maternal immune responses [1]. Due to its critical importance for maternal and fetal health, placental abnormalities are associated with up to 65% of fetal death cases [2].

Given its significance, any detected anomalies in the placenta or related structures—such as the umbilical cord and amniotic membranes—should be identified and subjected to histopathological examination [3]. While such examinations provide valuable insights into neonatal development and support future family planning [4], performing them after every delivery is costly and often impractical. As a result, they are typically reserved for cases where macroscopic analysis indicates a need, as determined by the attending obstetrician.

In recent years, artificial intelligence has emerged as a promising tool for medical image analysis, particularly through deep learning techniques [5]. Despite significant advances in prenatal imaging—such as ultrasound and magnetic resonance imaging (MRI) [1, 6, 7] - the application of AI to postnatal macroscopic placental analysis remains limited [8].

In this study, we propose a computer vision approach to automatically classify the presence or absence of macroscopic anomalies in human placentas, rather than classifying specific types of anomalies. To this end, we evaluate the performance of four convolutional neural network (CNN) architectures - ResNet-18, ResNet-34, EfficientNet-B0, and EfficientNet-B4 - using various hyperparameter configurations. Model performance is assessed based on four key metrics: accuracy, precision, recall, and F1 score.

2. Related work

This section explores existing solutions for classifying the presence of macroscopic anomalies in human placentas. Only studies published between 2019 and 2025 were considered, with a particular focus on the use of neural networks.

Among these studies, Y. Chen et al. [9] proposed a two-stage pipeline based on morphological characterization and feature analysis of the placenta. The first stage focuses on placenta segmentation, classification of the maternal and fetal sides, and localization of the umbilical cord insertion point. The second stage leverages transfer learning using various convolutional neural network (CNN) architectures (ResNet and YOLO), along with additional techniques, to identify a wide range of placental anomalies—including retained placenta, abnormal cord insertion, meconium, abruption, chorioamnionitis, irregular shape, hypercoiled cord, and true knots—achieving promising results across all classifications.

Z. Zhang et al. [10] introduced an approach based on multi-region saliency learning (MSL) for cross-domain segmentation of placental images, aiming to address the generalization challenges caused by morphological and lighting variations across datasets. The MSL strategy preserves critical semantic features when transferring knowledge between domains, enabling robust model adaptation. This technique demonstrated high accuracy in predicting inflammation associated with chorioamnionitis, even in the absence of annotations in the target domain, highlighting its clinical potential in data-scarce environments.

Y. Pan et al. [11] proposed the use of vision-and-language contrastive learning (VLC) to address data scarcity by incorporating pathology reports alongside placental images. The images were labeled with the following abnormalities: meconium, fetal inflammatory response (FIR), maternal inflammatory response (MIR), chorioamnionitis, and neonatal sepsis. Using ResNet50 and BERT architectures with additional neural network layers, the approach achieved higher AUC-ROC scores across all classes when compared to standalone ResNet50 and the ConVIRT method.

In a subsequent study, Y. Pan et al. [8] presented improvements to the VLC framework. First, they introduced report feature recomposition to prevent the encoder from suppressing relevant features during loss minimization, thereby allowing reports of varying lengths to be effectively processed. Second, they implemented distributional report feature recomposition to better compute similarity between embedding vectors. Coupled with architectures such as ResNet50, EfficientNet-B0, and MobileNetV3, this enhanced method achieved significantly better AUC-ROC scores across all classes, along with improved training and inference efficiency.

Silva et al. [12] proposed at the IEEE International Symposium on Computer-Based Medical Systems (CBMS) 2025 a study where three convolutional architectures were evaluated for placental segmentation: YOLO11, ResNet34 and EfficientNet-B0. The YOLO11 achieved an average AP50 of 72% and AP50-AP95 of 67.1%. The ResNet34 architecture reached a classification accuracy of 72.86%, with the mask segmentation reaching a Dice coefficient of 61.97% and an average IoU of 49.91%. The EfficientNet-B0 obtained an accuracy of 85.18%, but its mask segmentation showed an average Dice score of 31.55% and mean IoU of 24.70%.

3. Dataset

We utilized a dataset comprising 185 placental images, both standardized and non-standardized, collected from deliveries performed at Obstetric Centers and Maternity Hospitals. The images obtained at the University Hospital Polydoro Ernani de São Thiago, Federal University of Santa Catarina (HU-UFSC/EBSERH), were acquired following a standardized imaging protocol [12]. The dataset was categorized into two classes: altered placentas and non-altered placentas.

The altered class includes images exhibiting one or more of the following conditions: marginal insertion of the umbilical cord with radial blood vessel distribution, peripheral cord insertion with radial vasculature, irregular placental disc, oval-shaped disc, opaque placental membranes, placental lesions,

and greenish membranes - commonly associated with meconium. The number of images for each condition is shown in Table 1. Figure 1 shows one sample for each condition.

Table 1. Distribution of placental conditions in the dataset, including normal and altered cases

Alteration	Count	Proportion
None	154	83.2%
Marginal insertion of the umbilical cord with radial blood vessel distribution	3	1.6%
Peripheral cord insertion with radial vasculature	3	1.6%
Irregular placental disc	1	0.5%
Oval-shaped disc	2	1.1%
Opaque placental membranes	4	2.2%
Placental lesions	16	8.6%
Greenish membranes	2	1.1%
Total	**185**	**100%**

The dataset was split into training, validation, and test sets following a ratio of 0.8, 0.1, and 0.1, respectively. The distribution of cases across these sets is presented in Table 2.

Table 2. Distribution of placenta images by class across training, validation, and test sets

Class	Total	Train	Validation	Test
Unaltered	154	124	15	15
Altered	31	25	3	3
Total	**185**	**149**	**18**	**18**

3.1 Data augmentation

Data augmentation is a widely used technique in the training of machine learning models. It involves applying artificial transformations to input images in order to increase the diversity of the dataset without requiring the collection of new samples. These transformations may include operations such as rotations, translations, flipping, and changes in brightness and contrast, among others. They are applied stochastically during training. This approach significantly enhances the model's generalization ability, reducing the risk of overfitting and making the model more robust to real-world variations.

In this work, the experiments were divided into two scenarios: with and without data augmentation.

The scenario without data augmentation, hereafter referred to as NoDA, uses the original images without the application of any transformations before being processed by the network.

The scenario with data augmentation, referred to as DA, involves applying both geometric and photometric transformations to the images in order to enrich the training set. The following transformations were used in this study:

- Horizontal flip: flips the image horizontally with a 50% probability;
- Rotation: randomly rotates the image between -10° and +10°, applied with a 75% probability;
- Brightness and contrast adjustment: randomly alters the brightness and contrast by up to 10% each, with a 75% probability;
- Shear (tilt transformation): applies a shear of up to 20% to the image, with a 75% probability.

The results of the experiments are presented for both scenarios in order to assess the impact of data augmentation on the performance of the trained models.

4. Methods

4.1 Neural networks

We evaluated four convolutional neural network (CNN) models in total, based on two distinct architectures: ResNet-18, ResNet-34, EfficientNet-B0, and EfficientNet-B4. The ResNet architecture is well known for its use of residual blocks, which address the vanishing gradient problem [13], while the EfficientNet architecture employs compound scaling to balance computational efficiency with model accuracy [14].

ResNet-18 is a CNN model composed of 18 layers, including convolutional layers and residual blocks designed to mitigate the vanishing gradient issue. Each residual block consists of two convolutional layers, with a skip connection that adds the input of the block directly to the output of the second convolutional layer. This connection allows gradients to flow more easily through deeper networks. ResNet-34 follows the same residual block mechanism but extends the depth of the network to 34 layers, enabling the learning of more complex features.

EfficientNet-B0 is a lightweight CNN model that achieves high accuracy while maintaining excellent computational efficiency. It introduces a compound scaling method that uniformly scales the network's depth, width, and resolution using a predefined compound coefficient, instead of scaling each dimension arbitrarily. This design enables the network to achieve a favorable trade-off between performance and resource usage, reducing both the number of parameters and floating-point operations per second (FLOPs). EfficientNet-B4 applies the same compound scaling principles but with a larger network, increasing the number of layers and channels per layer to further improve accuracy.

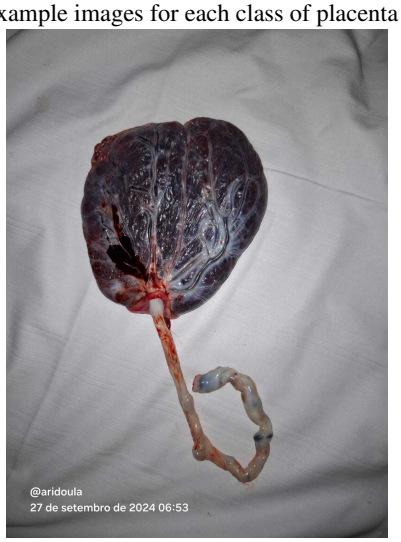
4.2 Hyperparameters

As previously mentioned, the dataset was randomly split into training, validation, and test sets with proportions of 0.8, 0.1, and 0.1, respectively.

To determine the learning rate, we used the *LR find* utility provided by the FastAI library [15]. This method performs a training cycle where the learning rate increases exponentially

Figure 1. Example images for each class of placenta alterations.

(a) No alteration

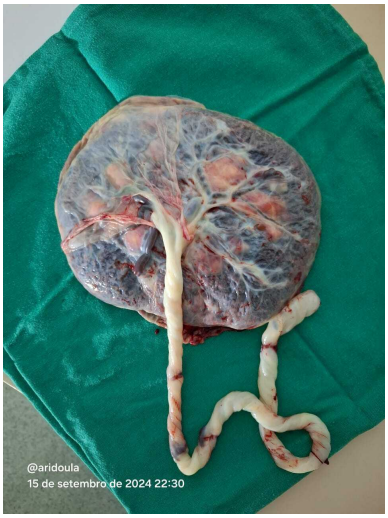


(b) Marginal cord insertion

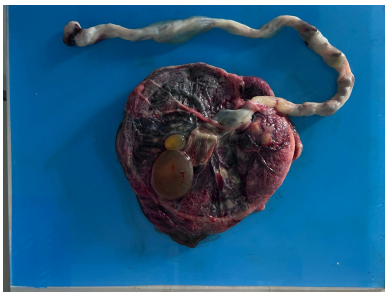
(c) Peripheral cord insertion

(d) Irregular disc

(e) Oval-shaped disc



(f) Opaque membranes



(g) Placental lesions

(h) Greenish membranes

at each iteration while monitoring the loss function. The goal is to identify the range in which the loss decreases most rapidly, thereby avoiding excessively high learning rates that could destabilize training. Based on this method, we adopted the library's automatic heuristic to select a value near the point of steepest descent or local minimum in the loss curve.

For batch size (BS) optimization, we applied a grid search strategy over the following values: BS = 8, 16, 32

To further enhance training, we adopted the One Cycle Policy, introduced by Leslie N. Smith (2018) [16]. This learning rate scheduling strategy involves cyclically varying both the learning rate and momentum during training—initially increasing the learning rate to a maximum and then gradually decreasing it to a very low value. This single-cycle schedule enables the model to explore a broader parameter space at the beginning, encouraging better generalization, and then fine-tune the learned parameters towards convergence. This policy has been shown to accelerate training, improve model performance, prevent overfitting, and reduce the need for extensive manual hyperparameter tuning.

Additionally, we employed the Adam (Adaptive Moment Estimation) optimizer to update model weights during training. Adam combines the benefits of two widely used methods: Momentum, which accumulates gradients from previous iterations to accelerate convergence, and RMSProp, which adapts the learning rate for each parameter based on a moving average of recent squared gradients. This combination makes Adam particularly effective for large-scale problems, offering fast and stable convergence.

The models were trained without fine-tuning. Preliminary results obtained using transfer learning alone indicated that the selected CNN architectures were capable of achieving up to 100% accuracy in classifying macroscopic placental anomalies. These results will be discussed in detail in the following sections.

4.3 Computational environment

The experiments were conducted using the Google Colab environment, with accelerated GPU support provided by an NVIDIA Tesla T4 graphics card.

5. Evaluations

We used four traditional classification metrics to evaluate the performance of our experiments: accuracy, recall (sensitivity), precision, and F1 score. These metrics provide insights into the model's ability to generalize learned knowledge to previously unseen samples, offering complementary perspectives on performance.

These metrics are derived from the following fundamental concepts:

- TP (True Positives): number of positive examples correctly classified
- TN (True Negatives): number of negative examples correctly classified
- FP (False Positives): number of negative examples incorrectly classified as positive
- FN (False Negatives): number of positive examples incorrectly classified as negative

The evaluation metrics are defined as follows:

- Accuracy (Equation 1): the proportion of correct predictions among all evaluated samples.
- Recall/Sensitivity (Equation 2): measures the model's ability to correctly identify positive samples, i.e., the fraction of true positives out of all actual positives.
- Precision (Equation 3): the proportion of true positives among all samples predicted as positive, reflecting the correctness of the model's positive predictions.
- F1 Score (Equation 4): the harmonic mean of precision and recall, providing a balanced measure that accounts for both false positives and false negatives.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

$$F1-Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

6. Results and discussion

Tables 3, 4 and 5 present the results obtained from the training with various hyperparameters, as described in Section 4.2.

The tables report the outcomes for each architecture and hyperparameter tuning strategy, including the use of data augmentation and variations in batch size. It can be observed that the results are remarkably similar across different architectures and hyperparameter settings, achieving an F1-score of 1.0 and accuracy values close or equal to 1.0 in all evaluated scenarios.

Figures 2, 3, 4 present the confusion matrixes of an aggregation of architectures and configurations with similar results.

Analyzing the individual metrics:

- Accuracy consistently approached 1.0, indicating that the models made almost no incorrect predictions across all test samples.
- F1-Score, as a harmonic mean of precision and recall, also reached 1.0, suggesting perfect balance between sensitivity and specificity in the classification task.

Table 3. Experiment results with the models trained with different data augmentation strategies with a batch size of 8.

		VALIDATION		TEST	
	Architecture	**NoDA**	**DA**	**NoDA**	**DA**
Accuracy	ResNet-18	0.944444	1.000000	1.000000	1.000000
	ResNet-34	1.000000	0.944444	0.944444	0.944444
	EfficientNet B0	1.000000	1.000000	0.888889	0.944444
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
Precision	ResNet-18	0.875000	1.000000	1.000000	1.000000
	ResNet-34	1.000000	0.968750	0.875000	0.875000
	EfficientNet B0	1.000000	1.000000	0.800000	0.875000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
Recall	ResNet-18	0.966667	1.000000	1.000000	1.000000
	ResNet-34	1.000000	0.833333	0.966667	0.966667
	EfficientNet B0	1.000000	1.000000	0.933333	0.966667
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
F1-Score	ResNet-18	0.911330	1.000000	1.000000	1.000000
	ResNet-34	1.000000	0.883871	0.911330	0.911330
	EfficientNet B0	1.000000	1.000000	0.839286	0.911330
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000

Table 4. Experiment results with the models trained with different data augmentation strategies with a batch size of 16.

		VALIDATION		TEST	
	Architecture	**NoDA**	**DA**	**NoDA**	**DA**
Accuracy	ResNet-18	1.000000	1.000000	0.944444	1.000000
	ResNet-34	0.944444	1.000000	1.000000	1.000000
	EfficientNet B0	1.000000	1.000000	0.944444	1.000000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
Precision	ResNet-18	1.000000	1.000000	0.875000	1.000000
	ResNet-34	0.968750	1.000000	1.000000	1.000000
	EfficientNet B0	1.000000	1.000000	0.875000	1.000000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
Recall	ResNet-18	1.000000	1.000000	0.966667	1.000000
	ResNet-34	0.833333	1.000000	1.000000	1.000000
	EfficientNet B0	1.000000	1.000000	0.966667	1.000000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
F1-Score	ResNet-18	1.000000	1.000000	0.911330	1.000000
	ResNet-34	0.883871	1.000000	1.000000	1.000000
	EfficientNet B0	1.000000	1.000000	0.911330	1.000000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000

Table 5. Experiment results with the models trained with different data augmentation strategies with a batch size of 32.

		VALIDATION		TEST	
	Architecture	**NoDA**	**DA**	**NoDA**	**DA**
Accuracy	ResNet-18	1.000000	1.000000	0.888889	1.000000
	ResNet-34	0.944444	1.000000	1.000000	0.944444
	EfficientNet B0	1.000000	1.000000	0.944444	1.000000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
Precision	ResNet-18	1.000000	1.000000	0.800000	1.000000
	ResNet-34	0.968750	1.000000	1.000000	0.875000
	EfficientNet B0	1.000000	1.000000	0.875000	1.000000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
Recall	ResNet-18	1.000000	1.000000	0.933333	1.000000
	ResNet-34	0.833333	1.000000	1.000000	0.966667
	EfficientNet B0	1.000000	1.000000	0.966667	1.000000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000
F1-Score	ResNet-18	1.000000	1.000000	0.839285	1.000000
	ResNet-34	0.883871	1.000000	1.000000	0.911330
	EfficientNet B0	1.000000	1.000000	0.911330	1.000000
	EfficientNet B4	1.000000	1.000000	1.000000	1.000000

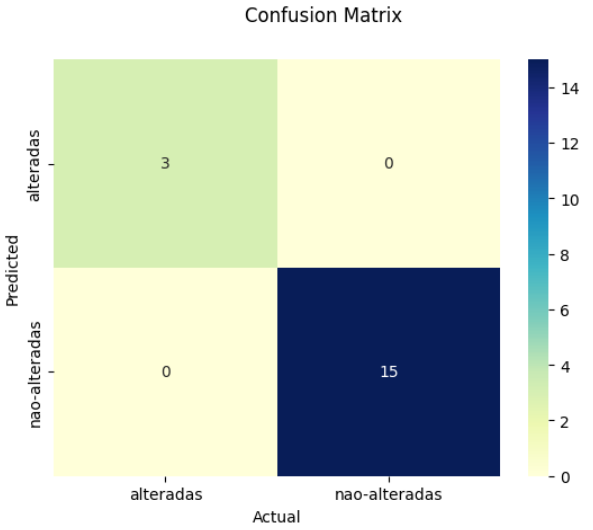


Figure 2. Confusion matrix for the configurations that reached an F1-Score of **1.000**. Architectures/Configurations: ResNet-18 (NoDA, BS=8), EfficientNet-B4 (NoDA, BS=8), ResNet-18 (DA, BS=8), EfficientNet-B4 (DA, BS=8), ResNet-34 (NoDA, BS=16), EfficientNet-B4 (NoDA, BS=16), ResNet-18 (DA, BS=16), ResNet-34 (DA, BS=16), EfficientNet-B0 (DA, BS=16), EfficientNet-B4 (DA, BS=16), ResNet-18 (NoDA, BS=32), EfficientNet-B4 (DA, BS=32), ResNet-18 (DA, BS=32), ResNet-34 (DA, BS=32), EfficientNet-B4 (DA, BS=32).

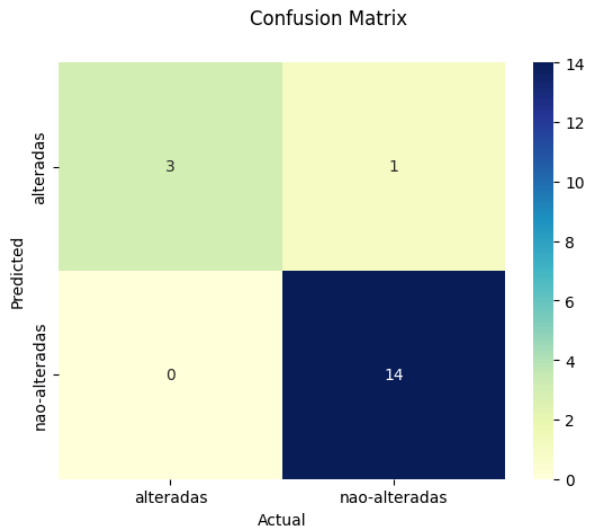


Figure 3. Confusion matrix for the configurations that reached an F1-Score of **0.884**. Architectures/Configurations: ResNet-34 (NoDA, BS=8), ResNet-34 (DA, BS=8), EfficientNet-B0 (DA, BS=8), ResNet-18 (NoDA, BS=16), EfficientNet-B0 (NoDA, BS=16), EfficientNet-B0 (NoDA, BS=32), ResNet-34 (DA, BS=32).

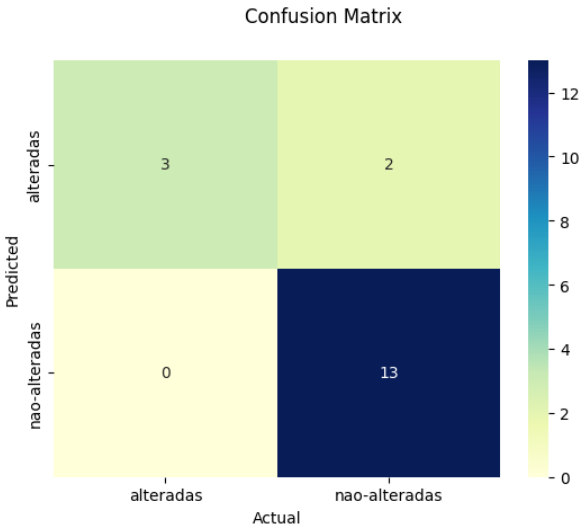


Figure 4. Confusion matrix for the configurations that reached an F1-Score of **0.839**. Architectures/Configurations: EfficientNet-B0 (NoDA, BS=32), ResNet-18 (NoDA, BS=32).

The main threat to this study's validity is the limited dataset size and significant class imbalance. Many images differ only by subtle variations in angle, lighting, or perspective, resulting in highly similar samples across the training, validation, and test sets. This homogeneity increases the risk of overfitting, where models learn to memorize specific patterns rather than develop generalizable representations. Consequently, the reported performance may overestimate the models' true predictive ability in more diverse clinical settings.

To mitigate this issue, we propose expanding the dataset with additional placental images showing similar macroscopic abnormalities but sourced from distinct clinical cases. This would allow for a more robust assessment of the models' generalization capabilities and support the refinement of the network architectures under more realistic and varied conditions. Such efforts are critical for enhancing the clinical relevance and reliability of the proposed method.

7. Conclusion

In this study, two neural network architectures were compared for the classification of placental images with and without macroscopic abnormalities. Pre-trained models were employed along with a series of hyperparameter configurations and data augmentation techniques to determine the best-performing setup for each architecture. The results demonstrate that convolutional neural networks are effective in classifying placentas with various morphological alterations, achieving F1-score and accuracy values close to 1.0.

Despite these promising results, we acknowledge the limitations in the generalization potential of the models given the current dataset. Therefore, further research is required to evaluate this capability more rigorously and to develop a more comprehensive dataset. Expanding the dataset will be essential to enhance the generalizability and robustness of the proposed models in real-world clinical applications.

Author contributions

- Conceptualization: A.W., B.Q.T.S., B.S.L., and S.K.
- Methodology: A.W., B.Q.T.S., B.S.L., and S.K.
- Development: A.W., B.S.L., and S.K.
- Discussion of results: A.W., B.S.L., and S.K.
- Writing: B.S.L. and S.K
- Review & Editing: A.W., B.Q.T.S., B.S.L., and S.K.
- Supervision: A.W. and B.Q.T.S.

References

[1] ALANSARY, A. et al. Fast fully automatic segmentation of the human placenta from motion corrupted mri. In: *Congresso Internacional de Processamento de Imagens Médicas e Intervenção Assistida por Computador*. [S.l.: s.n.], 2016. p. 589–597.

[2] LANGSTON, C. et al. Practice guideline for examination of the placenta: developed by the placental pathology practice guideline development task force of the college of american pathologists. *Arch Pathol Lab Med*, v. 121, n. 5, p. 449–476, 1997.

[3] MITTAL, N. et al. A practical guide to placental examination for forensic pathologists. *Forensic Science, Medicine and Pathology*, v. 16, p. 295–312, 2020. Disponível em: ⟨https://doi.org/10.1007/s12024-019-00214-2⟩.

[4] KHONG, T. Y. et al. Sampling and definitions of placental lesions: Amsterdam placental workshop group consensus statement. *Arch Pathol Lab Med*, v. 140, n. 7, p. 698–713, 2016.

[5] LITJENS, G. et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, Elsevier, v. 42, p. 60–88, 2017.

[6] MALATHI, G.; SHANTHI, V. Statistical measurement of ultrasound placenta images complicated by gestational diabetes mellitus using segmentation approach. *J Inform Hiding Multimedia Signal Process*, p. 332–343, 2011.

[7] LOONEY, P. et al. Automatic 3d ultrasound segmentation of the first trimester placenta using deep learning. In: *IEEE International Symposium on Biomedical Imaging*. [S.l.: s.n.], 2017. p. 279–282.

[8] PAN, Y. et al. Enhancing automatic placenta analysis through distributional feature recomposition in vision-language contrastive learning. In: GREENSPAN, H. et al. (Ed.). *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*. Cham: Springer Nature Switzerland, 2023. p. 116–126. ISBN 978-3-031-43987-2.

[9] CHEN, Y. et al. Ai-plax: Ai-based placental assessment and examination using photos. *Computerized Medical Imaging and Graphics*, v. 84, p. 101744, 2020. ISSN 0895-6111. Disponível em: ⟨https://www.sciencedirect.com/science/article/pii/S0895611120300471⟩.

[10] ZHANG, Z. et al. Multi-region saliency-aware learning for cross-domain placenta image segmentation. *Pattern Recognition Letters*, v. 140, p. 165–171, 2020. ISSN 0167-8655. Disponível em: ⟨https://www.sciencedirect.com/science/article/pii/S0167865520303780⟩.

[11] PAN, Y. et al. Vision-language contrastive learning approach to robust automatic placenta analysis using photographic images. In: WANG, L. et al. (Ed.). *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. Cham: Springer Nature Switzerland, 2022. p. 707–716. ISBN 978-3-031-16437-8.

[12] SILVA, B. Q. T. da et al. Towards automated placental screening: Instance segmentation in clinical images. In: *Proceedings of the 38th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*. [s.n.], 2025. To appear. Disponível em: ⟨https://2025.cbms-conference.org/⟩.

[13] HE, K. et al. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 770–778.

[14] TAN, M.; LE, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. In: *Proceedings of the 36th International Conference on Machine Learning (ICML)*. [S.l.]: PMLR, 2019. (Proceedings of Machine Learning Research, v. 97), p. 6105–6114.

[15] HOWARD, J.; GUGGER, S. fastai: A layered api for deep learning. *Information*, v. 11, n. 2, p. 108, 2020.

[16] SMITH, L. N.; TOPIN, N. Super-convergence: Very fast training of neural networks using large learning rates. *arXiv preprint arXiv:1708.07120*, 2017. Disponível em: ⟨https://arxiv.org/abs/1708.07120⟩.