



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CAMPUS FLORIANÓPOLIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Beatriz Silva Lopes

**Visão Computacional com Deep Learning como Metodologia de Suporte à  
Triagem Automatizada de Alterações Macroscópicas Placentárias e Solicitação  
de Exame Anatomopatológico**

Florianópolis  
2026

Beatriz Silva Lopes

**Visão Computacional com Deep Learning como Metodologia de Suporte à  
Triagem Automatizada de Alterações Macroscópicas Placentárias e Solicitação  
de Exame Anatomopatológico**

Dissertação submetida ao Programa de Pós-Graduação  
em Ciência da Computação da Universidade Federal  
de Santa Catarina para a obtenção do título de Mes-  
tra em Ciência da Computação.

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

Florianópolis  
2026

Ficha catalográfica gerada por meio de sistema automatizado gerenciado pela BU/UFSC.  
Dados inseridos pelo próprio autor.

Silva Lopes, Beatriz

Visão Computacional com Deep Learning como Metodologia de Suporte à Triagem Automatizada de Alterações Macroscópicas Placentárias e Solicitação de Exame Anatomopatológico / Beatriz Silva Lopes ; orientador, Aldo von Wangenheim, 2026.

111 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Ciência da Computação, Florianópolis, 2026.

Inclui referências.

1. Ciência da Computação. 2. Placenta. 3. Visão Computacional. 4. Rede Neural Convolucional. 5. Segmentação de Imagem. I. Wangenheim, Aldo von. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Ciência da Computação. III. Título.

Beatriz Silva Lopes

**Visão Computacional com Deep Learning como Metodologia de Suporte à  
Triagem Automatizada de Alterações Macroscópicas Placentárias e Solicitação  
de Exame Anatomopatológico**

O presente trabalho em nível de Mestrado foi avaliado e aprovado, em 09 de dezembro de 2025, pela banca examinadora composta pelos seguintes membros:

Prof. Alex Sandro Roschildt Pinto, Dr.  
Universidade Federal de Santa Catarina

Prof. Antonio Carlos Sobieranski, Dr.  
Universidade Federal de Santa Catarina

Prof. Douglas Dyllon Jeronimo de Macedo, Dr.  
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestra em Ciência da Computação.

---

Coordenação do Programa de  
Pós-Graduação

---

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

Florianópolis, 2026.

Dedico este trabalho aos meus pais, Neilson e Regiane, meus irmãos, Amanda, Viviane e Vitor, minha avó, Áurea e, meu companheiro, Cleber, por todo o amor, apoio e compreensão que tornaram este caminho mais leve e vibrante. Com imensa saudade, ao meu fiel amigo Bartholomeu (*in memoriam*). Sua lealdade e amor incondicional nas madrugadas de estudo foram paz e eterna inspiração.

## AGRADECIMENTOS

A conclusão desta etapa da jornada acadêmica, simbolizada pela defesa desta Dissertação, é um marco que reflete o esforço e o apoio incondicional de inúmeras pessoas e instituições, às quais dedico minha mais profunda e sincera gratidão.

Aos meus amados pais, irmãos, sobrinhos e avós, meu eterno e mais sincero agradecimento por serem meu porto seguro, pelo apoio, pelo investimento constante e por estarem sempre ao meu lado, mesmo distante, acompanhando cada nova decisão e passo.

Ao meu namorado e enteados, dedico meu profundo reconhecimento por terem caminhado ao meu lado do começo ao fim desta jornada, sendo um pilar fundamental ao longo de todos os momentos, fornecendo apoio contínuo, incentivo e muito amor.

Com imensa saudade, agradeço também o meu fiel amigo Bartholomeu (*in memoriam*), pela luz, alegria e amor incondicional que me acompanharam nas longas madrugadas de estudo.

Aos meus amigos e colegas, um obrigado pela compreensão e paciência durante os meses de dedicação exclusiva.

Agradeço imensamente ao meu orientador, Prof. Dr. rer. nat. Aldo von Wangenheim, pela confiança depositada, pela excelência e pela generosidade em compartilhar sua vasta experiência e conhecimento, orientando este trabalho com rigor científico.

À Dra. Bibiana Tiellet, minha amiga e parceira de pesquisa, que me introduziu ao fascinante mundo da medicina.

Um agradecimento especial aos membros da banca avaliadora, Prof. Dr. Douglas Dyllon Jeronimo de Macedo, Prof. Dr. Antonio Carlos Sobieranski e Prof. Dr. Alex Sandro Roschildt Pinto, por aceitarem o convite para participar na defesa final.

Aos profissionais de saúde que generosamente apoiaram o complexo processo de captura de imagens, e em especial, a todas as mulheres no pós-parto que, tornaram este estudo possível e viável.

Agradeço à Universidade Federal de Santa Catarina (UFSC) e ao Programa de Pós-Graduação em Ciência da Computação (PPGCC), pela oportunidade de evoluir em minha caminhada acadêmica.

Por fim, expresso meu reconhecimento institucional ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de pós-graduação concedida, e à Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC), que forneceu os recursos computacionais essenciais para a execução deste trabalho (processo nº 2024TR000090).

*"A mente que se abre a uma nova ideia  
jamais voltará ao seu tamanho original."  
(EINSTEIN, Albert)*

## RESUMO

O exame anatomopatológico da placenta é crucial para elucidar desfechos perinatais adversos, mas sua solicitação sistemática é inviável devido aos altos custos e demanda operacional. A triagem inicial, baseada na avaliação macroscópica no momento do parto, é fundamental, mas sofre com a subjetividade e a não-padronização. Diante dessa limitação, este trabalho propôs e validou um pipeline computacional de Deep Learning (DL) para a análise automatizada de imagens macroscópicas de placentas, visando fornecer um sistema de suporte à decisão para a triagem clínica. A metodologia incluiu o desenvolvimento e documentação de um Protocolo de Captura de Imagens (PCI) padronizado para a sala de parto, o que permitiu a construção de um dataset proprietário, abrangente e anotado por especialistas em nove classes morfológicas. Foram exploradas e comparadas quatorze arquiteturas de DL, focadas em segmentação semântica e segmentação por instância, para estabelecer um benchmark de desempenho. Os resultados mostraram que a arquitetura U-Net com backbone ResNet34 obteve o melhor desempenho global em segmentação semântica, com Dice Coefficient de 80,71% e Intersection over Union (IoU) de 70,01%. A U-Net com EfficientNet-B7 alcançou a maior acurácia global (91,33%) e o menor tempo de execução entre as redes U-Net avaliadas. No que concerne à segmentação por instância, o modelo YOLOv11 obteve um mAP50 de 69,5% e mAP50-95 de 52,8%. Conclui-se que os modelos demonstram viabilidade técnica e bom desempenho quantitativo, avançando em direção à robustez necessária para apoiar a triagem automatizada, embora a generalização e robustez plena dependam da expansão e diversificação do conjunto de dados. O estudo estabelece uma base sólida para o desenvolvimento futuro de ferramentas assistivas que podem otimizar a alocação de recursos hospitalares e contribuir diretamente para a saúde materno-fetal.

**Palavras-chave:** Placenta. Patologia. Rede Neural Convolutacional. Segmentação de Imagem. Exame pós-parto. Visão Computacional.

## ABSTRACT

Histopathological examination of the placenta is crucial for clarifying adverse perinatal outcomes, but its systematic request is unfeasible due to high costs and operational demands. Initial screening, based on macroscopic evaluation at the time of delivery, is fundamental but suffers from subjectivity and non-standardization. Given this limitation, this work proposed and validated a Deep Learning (DL) computational pipeline for the automated analysis of macroscopic placental images, aiming to provide a clinical triage decision support system. The methodology included the development and documentation of a standardized Protocol for Image Capture (PCI) in the delivery room, which allowed for the construction of a proprietary, comprehensive dataset annotated by specialists into nine morphological classes. Fourteen DL architectures were explored and compared, focusing on semantic segmentation and instance segmentation, to establish a performance benchmark. The results showed that the U-Net architecture with the ResNet34 backbone achieved the best overall performance in semantic segmentation, with a Dice Coefficient of 80.71% and an Intersection over Union (IoU) of 70.01%. The U-Net with EfficientNet-B7 achieved the highest global accuracy (91.33%) and the shortest execution time among the evaluated U-Net networks. Regarding instance segmentation, the YOLOv11 model achieved an mAP50 of 69.5% and mAP50-95 of 52.8%. It is concluded that the models demonstrate technical feasibility and good quantitative performance, advancing toward the necessary robustness to support automated triage, although full generalization and robustness depend on the expansion and diversification of the dataset. The study establishes a solid foundation for the future development of assistive tools that can optimize the allocation of hospital resources and directly contribute to maternal and fetal health.

**Keywords:** Placenta. Pathology. Deep Learning. Image Segmentation. Postpartum Screening. Computer Vision.

## LISTA DE FIGURAS

|   |    |
|---|----|
| Figura 1 – Fluxo de trabalho proposto para a triagem placentária automatizada.  | 18 |
| Figura 2 – Metodologia adotada.   | 19 |
| Figura 3 – Fases da Pesquisa.   | 20 |
| Figura 4 – Análise macroscópica da placenta através de representações esquemáticas (A) e imagens reais (B). Lado fetal brilhante com vasos sanguíneos dispostos em padrão arboriforme e cordão umbilical conectado ao centro do disco. Lado materno intacto e discoide, com textura esponjosa cor de vinho coberta por tecido uterino (decídua) e contendo alguns coágulos. | 23 |
| Figura 5 – Hierarquia e Relação entre os Principais Campos da Inteligência Artificial (IA).   | 25 |
| Figura 6 – Fluxo base Machine Learning (ML) aplicado a classificação de imagem.   | 25 |
| Figura 7 – Fluxo base DL aplicado a classificação de imagem.  | 26 |
| Figura 8 – Representação do neurônio biológico.   | 27 |
| Figura 9 – Representação Redes Neurais Artificiais (RNA).   | 28 |
| Figura 10 – Representação Redes Neurais Convolucionais (CNN).   | 30 |
| Figura 11 – Arquitetura do PlacentaNet.   | 33 |
| Figura 12 – Diagrama esquemático do framework proposto de dois estágios para avaliação e exame automatizados da placenta utilizando fotos.  | 34 |
| Figura 13 – Pipeline MSL (Aprendizado de Saliência Multi-regional) proposto.  | 34 |
| Figura 14 – Pipeline VLC (Aprendizado Contrastivo de Visão e Linguagem) proposto.   | 35 |
| Figura 15 – Método avançada de VLC proposto para analisar automaticamente imagens fotográficas de placentas.  | 35 |
| Figura 16 – Fluxo de Execução do PCI.   | 38 |
| Figura 17 – Exemplo de Imagem de Placenta Padronizada (segundo o PCI) e Imagem Não Padronizada (sem a aplicação do PCI).  | 39 |
| Figura 18 – Exemplos de imagens sintéticas obtidas por <i>data augmentation</i> , simulando variações comuns em capturas com dispositivos móveis.   | 40 |
| Figura 19 – Exemplo de anotação no Roboflow.  | 41 |
| Figura 20 – Evolução das métricas de treinamento e validação da arquitetura U-Net ResNet34.   | 47 |
| Figura 21 – Evolução das métricas de treinamento e validação da arquitetura U-Net EfficientNet-B7.  | 47 |
| Figura 22 – Exemplos comparativos entre imagem de entrada, verdade fundamental e predição obtida pela ResNet34.   | 48 |

|  |    |
|--|----|
| Figura 23 – Exemplos comparativos entre imagem de entrada, verdade fundamental e predição obtida pela EfficientNet-B7. . . . . | 49 |
| Figura 24 – Matriz de confusão gerada para ResNet34. . . . .   | 50 |
| Figura 25 – Matriz de confusão gerada para EfficientNet-B7. . . . .  | 50 |
| Figura 26 – Exemplos de casos mais desafiadores para a ResNet34. . . . .   | 51 |
| Figura 27 – Métricas por classe (detecção e segmentação por instância) obtidas com a YOLO. . . . .                             | 53 |
| Figura 28 – Matriz de confusão normalizada do modelo YOLOv11 nas tarefas de detecção e segmentação. . . . .                    | 53 |
| Figura 29 – Exemplos de anotações manuais e predições geradas pela YOLO. .   | 54 |

## LISTA DE QUADROS

|  |    |
|--|----|
| Quadro 1 – Distribuição da quantidade de instâncias rotuladas por classe morfológica utilizada no treinamento. . . . . | 40 |
|--|----|

## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 1 – Resumo da definição do protocolo utilizado na busca. . . . .                                     | 32 |
| Tabela 2 – Resultados e Processo de Filtragem da Busca. . . . .   | 33 |
| Tabela 3 – Configuração de treinamento para cada arquitetura avaliada. . . . .                              | 44 |
| Tabela 4 – Principais métricas de desempenho das arquiteturas U-Net (ResNet e EfficientNet). . . . .        | 46 |
| Tabela 5 – Métricas de detecção (box) e segmentação por instância (mask) obtidas com o modelo YOLO. . . . . | 52 |

## LISTA DE ABREVIATURAS E SIGLAS

|      |                                   |
|------|-----------------------------------|
| BD   | Banco de Dados                    |
| CNN  | Redes Neurais Convolucionais      |
| DL   | Deep Learning                     |
| IA   | Inteligência Artificial           |
| IoU  | Intersection over Union           |
| ML   | Machine Learning                  |
| PCI  | Protocolo de Captura de Imagens   |
| RNA  | Redes Neurais Artificiais         |
| RSL  | Revisão Sistemática da Literatura |
| VC   | Visão Computacional               |
| YOLO | You Only Live Once                |

## SUMÁRIO

|              |   |           |
|--------------|---|-----------|
| <b>1</b>     | <b>INTRODUÇÃO</b>                                     | <b>16</b> |
| 1.1          | MOTIVAÇÃO E JUSTIFICATIVA                             | 17        |
| 1.2          | HIPÓTESE  | 17        |
| 1.3          | OBJETIVOS   | 18        |
| <b>1.3.1</b> | <b>Objetivo Geral</b>                                 | <b>18</b> |
| <b>1.3.2</b> | <b>Objetivos Específicos</b>                          | <b>18</b> |
| 1.4          | METODOLOGIA   | 19        |
| 1.5          | ESTRUTURA DA DISSERTAÇÃO                              | 21        |
| <b>2</b>     | <b>FUNDAMENTAÇÃO TEÓRICA</b>                          | <b>22</b> |
| 2.1          | PLACENTA  | 22        |
| <b>2.1.1</b> | <b>Análise Macroscópica e Alterações Placentárias</b> | <b>23</b> |
| <b>2.1.2</b> | <b>Exame Anatomopatológico</b>                        | <b>24</b> |
| 2.2          | INTELIGÊNCIA ARTIFICIAL (IA)                          | 24        |
| <b>2.2.1</b> | <b>Machine Learning (ML)</b>                          | <b>25</b> |
| <b>2.2.2</b> | <b>Deep Learning (DL)</b>                             | <b>26</b> |
| <b>2.2.3</b> | <b>Redes Neurais Artificiais (RNA)</b>                | <b>27</b> |
| <b>2.2.4</b> | <b>Redes Neurais Convolucionais (CNN)</b>             | <b>28</b> |
| 2.3          | VISÃO COMPUTACIONAL (Visão Computacional (VC))        | 30        |
| <b>3</b>     | <b>TRABALHOS CORRELATOS</b>                           | <b>32</b> |
| 3.1          | DEFINIÇÃO E EXECUÇÃO DA BUSCA                         | 32        |
| 3.2          | SÍNTESE DOS RESULTADOS                                | 33        |
| 3.3          | DISCUSSÃO E POSICIONAMENTO DA PESQUISA                | 35        |
| <b>4</b>     | <b>DESENVOLVIMENTO</b>                                | <b>37</b> |
| 4.1          | MATERIAIS E AQUISIÇÃO DE DADOS                        | 37        |
| <b>4.1.1</b> | <b>Protocolo de Captura de Imagens (PCI)</b>          | <b>37</b> |
| 4.1.1.1      | Detalhamento de Execução do PCI                       | 37        |
| <b>4.1.2</b> | <b>Conjunto de Imagens (Dataset)</b>                  | <b>39</b> |
| 4.1.2.1      | Processamento e Classes                               | 39        |
| <b>4.1.3</b> | <b>Processo de Anotação e Exportação das Imagens</b>  | <b>40</b> |
| 4.2          | CONFIGURAÇÃO DO AMBIENTE                              | 41        |
| <b>4.2.1</b> | <b>Ambiente de Desenvolvimento</b>                    | <b>41</b> |
| <b>4.2.2</b> | <b>Arquiteturas Implementadas</b>                     | <b>42</b> |
| <b>4.2.3</b> | <b>Configurações e Estratégias de Treinamento</b>     | <b>43</b> |
| 4.3          | MÉTRICAS DE AVALIAÇÃO DE DESEMPENHO                   | 44        |
| <b>5</b>     | <b>RESULTADOS</b>                                     | <b>46</b> |
| 5.1          | DESEMPENHO GLOBAL                                     | 46        |
| <b>5.1.1</b> | <b>Métricas de Segmentação Semântica</b>              | <b>46</b> |

|       |   |            |
|-------|---|------------|
| 5.1.2 | <b>Métricas de Segmentação por Instância . . . . .</b>  | <b>51</b>  |
| 6     | <b>CONSIDERAÇÕES FINAIS . . . . .</b>   | <b>55</b>  |
| 6.1   | <b>DISCUSSÃO . . . . .</b>  | <b>55</b>  |
| 6.1.1 | <b>Ameaças à Validade . . . . .</b>   | <b>56</b>  |
| 6.2   | <b>CONCLUSÃO . . . . .</b>  | <b>57</b>  |
| 6.2.1 | <b>Trabalhos Futuros . . . . .</b>  | <b>58</b>  |
|       | <b>Referências . . . . .</b>  | <b>60</b>  |
|       | <b>APÊNDICE A – SYSTEMATIC LITERATURE REVIEW ON COM-<br/>PUTER VISION AND DEEP LEARNING FOR AU-<br/>TOMATED SCREENING OF MACROSCOPIC PLA-<br/>CENTAL ALTERATION . . . . .</b> | <b>64</b>  |
|       | <b>APÊNDICE B – ADVANCING AUTOMATED PLACENTAL SCREE-<br/>NING: DEEP LEARNING FOR MULTICLASS SEG-<br/>MENTATION IN POSTPARTUM IMAGE . . . . .</b>                              | <b>94</b>  |
|       | <b>ANEXO A – TOWARDS AUTOMATED PLACENTAL SCREENING:<br/>INSTANCE SEGMENTATION IN CLINICAL IMAGES .</b>  | <b>104</b> |

## 1 INTRODUÇÃO

A placenta é um órgão transitório fundamental para a troca de nutrientes, oxigênio e metabólitos entre mãe e feto, além de atuar na produção hormonal e no desenvolvimento da barreira imunológica fetal (ROBERTS, 2008). Alterações na morfologia placentária podem estar associadas a complicações gestacionais e desfechos perinatais adversos, sendo responsáveis por até 65% das causas de óbito fetal (PTACEK *et al.*, 2014). A análise da placenta, portanto, é um procedimento essencial para a investigação de eventos patológicos ocorridos durante a gestação (KHONG *et al.*, 2016).

A avaliação macroscópica da placenta permite a identificação de alterações estruturais do disco placentário e do cordão umbilical, bem como anomalias na coloração e integridade das membranas amnióticas. Em muitos casos, essa análise é determinante para a solicitação do exame anatomopatológico, realizado por patologistas mediante indicação do obstetra assistente (SILVA, T. *et al.*, 2016). No entanto, a solicitação sistemática do exame para todas as placentas é inviável devido a seu alto custo e demanda operacional. Assim, a decisão de envio para análise laboratorial é feita com base em critérios clínicos e características morfológicas observadas no momento do parto. No entanto, a subjetividade na avaliação macroscópica e a variação entre protocolos hospitalares podem impactar a detecção precoce de patologias placentárias (SILVA, T. *et al.*, 2016).

O avanço da Inteligência Artificial (IA) na área médica tem impulsionado o desenvolvimento de ferramentas computacionais para a análise de imagens médicas, incluindo exames pré-natais como ultrassonografia e ressonância magnética, e a avaliação tecidual por cortes histológicos. No entanto, no contexto da análise macroscópica pós-parto, observa-se uma lacuna na literatura quanto ao desenvolvimento de métodos automatizados para a identificação de padrões morfológicos em registros fotográficos de placentas.

Neste cenário de restrição, emerge a necessidade de ferramentas de triagem eficientes, rápidas e de baixo custo que possam auxiliar na seleção dos espécimes que, de fato, necessitam de investigação aprofundada. A IA, e em particular o Deep Learning (DL), tem se consolidado como uma metodologia promissora para a automação de processos de análise de imagens médicas (LE CUN; BENGIO; HINTON, 2015). A aplicação dessas técnicas à Visão Computacional (VC) na análise de imagens macroscópicas da placenta pós-parto representa uma oportunidade tecnológica para criar um sistema de suporte à decisão clínica no ponto de atendimento.

## 1.1 MOTIVAÇÃO E JUSTIFICATIVA

Embora o potencial da VC em auxiliar a Placentologia seja reconhecido, sua aplicação prática na análise macroscópica pós-parto ainda enfrenta desafios críticos. A Revisão Sistemática da Literatura (RSL) (LOPES; SILVA, B. Q. T.; WANGENHEIM, A. von, 2025b), anexada integralmente no Apêndice A revelou uma notável lacuna de pesquisa no estado da arte: a maior parte dos estudos concentra-se em imagens pré-natais (ultrassom) ou em fotografias de alta qualidade obtidas em ambientes controlados (laboratório de patologia).

O principal obstáculo reside na ausência de robustez e adaptabilidade em metodologias existentes, que falham ao processar a variabilidade e a não-padronização inerente a imagens capturadas em condições clínicas adversas (e.g., iluminação inconsistente, artefatos e enquadramento não padronizado na sala de parto). A variabilidade inerente à aquisição de imagens no ambiente clínico – incluindo a iluminação inconsistente e a falta de enquadramento padronizado – compromete a qualidade dos dados e torna o processo de anotação por especialistas complexo e demorado.

Com o objetivo de mitigar esse déficit, desenvolvemos um Protocolo de Captura de Imagem (PCI) padronizado para a placenta ainda na sala de parto (LOPES; SILVA, B. Q. T.; WANGENHEIM, A. von, 2025a), contemplando parâmetros como contraste de fundo, distância da câmera, iluminação e enquadramento. A partir desse protocolo, foi constituído um Banco de Dados (BD) próprio com imagens reais anotadas por patologistas, representando diferentes padrões morfológicos. A Figura 1 ilustra o fluxo de trabalho completo e a integração do pipeline proposto para a triagem placentária automatizada no ambiente clínico.

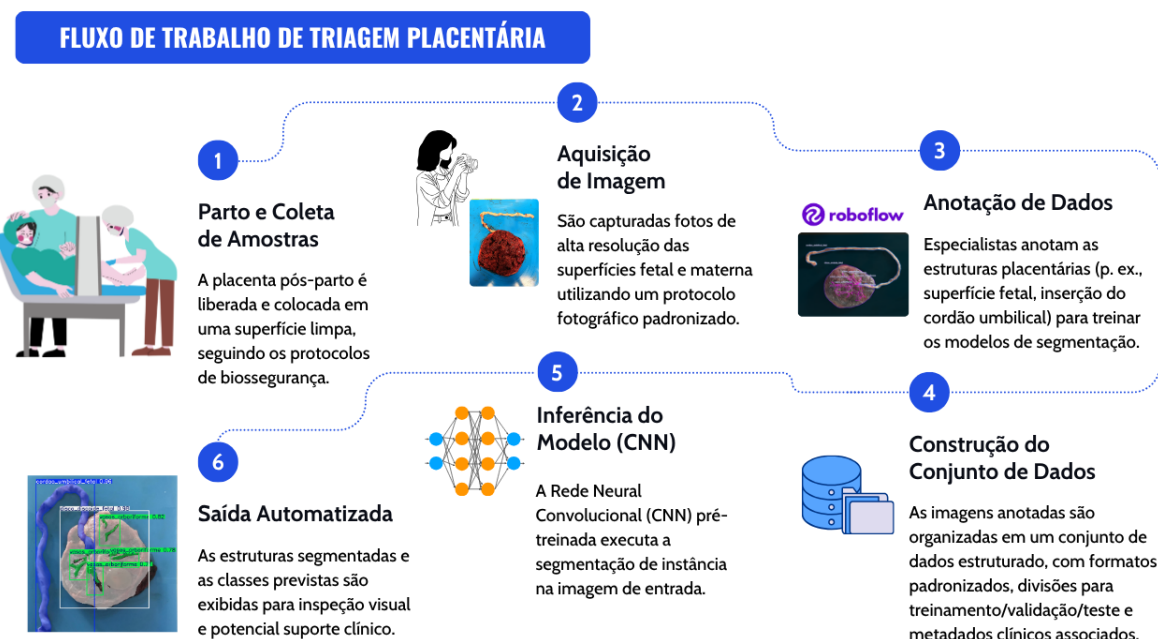
Diante disso, a motivação central deste trabalho é demonstrar a viabilidade técnica da triagem placentária automatizada no ambiente clínico real. A justificativa deste trabalho reside na urgência de transpor a barreira metodológica entre os avanços da Ciência da Computação e a aplicação clínica, fornecendo uma ferramenta que otimiza a alocação de recursos hospitalares e contribui diretamente para a saúde materno-fetal mediante a triagem objetiva.

## 1.2 HIPÓTESE

A avaliação macroscópica da placenta é um procedimento clínico fundamental, porém, a subjetividade inerente à análise manual e a ausência de protocolos padronizados resultam em inconsistência na triagem e em subaproveitamento dos recursos de anatomopatologia. Embora o potencial da VC tenha sido demonstrado, as abordagens existentes, frequentemente testadas em condições controladas, atingem um limite de generalização e robustez frente à variabilidade das imagens clínicas reais.

Neste contexto de restrição e desafio, a pergunta de pesquisa central que este

Figura 1 – Fluxo de trabalho proposto para a triagem placentária automatizada.



Fonte: Elaborado pela autora (2025).

trabalho se propõe a responder é: É possível desenvolver e validar um modelo computacional de DL, baseado em técnicas de segmentação, que ofereça um *benchmark* de precisão e uma robustez suficiente para suportar a decisão clínica na triagem automatizada de alterações macroscópicas placentárias em ambientes clínicos reais?

## 1.3 OBJETIVOS

### 1.3.1 Objetivo Geral

Desenvolver um pipeline computacional de DL para a análise automatizada de imagens macroscópicas de placentas, explorando e comparando abordagens de segmentação semântica e segmentação de instância, com o intuito de estabelecer um benchmark de desempenho e avaliar a viabilidade operacional da solução em um contexto clínico real.

### 1.3.2 Objetivos Específicos

De maneira específica, este trabalho busca:

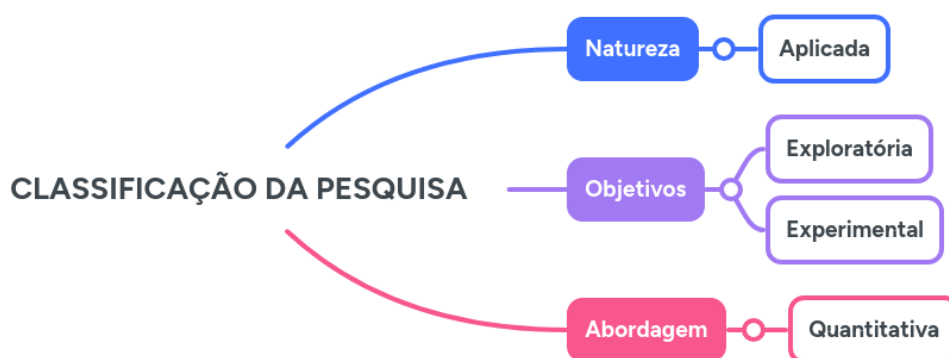
- Consolidar o estado da arte através da RSL para fundamentar a lacuna de pesquisa e a proposta metodológica.
- Desenvolver e documentar um PCI placentárias padronizado, garantindo a obtenção de dados de alta qualidade e otimizando o processo de anotação clínica.

- Constituir e anotar um BD abrangente de imagens macroscópicas de placentas, coletadas sob diferentes condições (imagens não padronizadas e imagens obtidas conforme o protocolo), incluindo a anotação multiclasse por especialistas.
- Implementar e experimentar múltiplas arquiteturas de DL.
- Avaliar o desempenho dos modelos computacionais obtidos e estabelecer um *benchmark* para a análise morfológica placentária, demonstrando a viabilidade técnica da solução.
- Recomendar as arquiteturas mais promissoras e propor as diretrizes formais para a validação clínica (Trabalhos Futuros).

#### 1.4 METODOLOGIA

Conforme a taxonomia estabelecida por Antonio C. Gil (GIL, 2002), o presente trabalho é classificado quanto à sua natureza, objetivos e abordagem. A Figura 2 mostra os tipos de pesquisas que serão utilizados neste trabalho.

Figura 2 – Metodologia adotada.



Fonte: Elaborado pela autora (2025).

##### **Quanto à natureza (finalidade):**

A pesquisa é classificada como Aplicada. Seu objetivo é gerar conhecimento que será aplicado diretamente na solução de um problema concreto e de interesse prático: o desenvolvimento de um pipeline computacional que sirva de apoio à decisão clínica na triagem placentária, otimizando recursos hospitalares.

##### **Quanto aos objetivos (meios):**

O estudo combina duas classificações distintas:

- **Pesquisa Exploratória:** Na fase inicial de RSL e no desenvolvimento do PCI, o estudo é Exploratório. Esta abordagem foi necessária devido à lacuna de conhecimento identificada na literatura sobre a aplicação de DL em imagens macroscópicas placentárias obtidas em ambientes clínicos reais, visando familiarizar-se com o problema e levantar hipóteses de trabalho.
- **Pesquisa Explicativa (ou Experimental):** Na fase de experimentação com DL, o trabalho assume um caráter Explicativo. O objetivo é identificar os fatores que influenciam o desempenho dos modelos. Por meio do controle e da manipulação das variáveis (as diferentes arquiteturas de Redes Neurais Convolucionais (CNN)), busca-se determinar e explicar qual modelo estabelece o *benchmark* de desempenho mais elevado para a triagem.

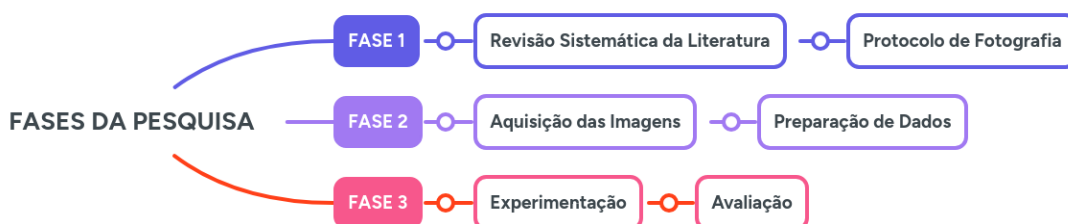
### Quanto à abordagem:

A pesquisa adota uma perspectiva predominantemente Quantitativa. A validação da hipótese central da pesquisa é baseada na coleta e análise de dados numéricos (métricas como Dice Score, IoU, mAP e etc), utilizando o tratamento estatístico para comparar o desempenho das arquiteturas e estabelecer o *benchmark* de forma objetiva.

### Pipeline de Execução:

A pesquisa é executada por meio de um pipeline estruturado em três fases principais, conforme ilustrado na Figura 3.

Figura 3 – Fases da Pesquisa.



Fonte: Elaborado pela autora (2025).

- **Fase I - Revisão e Protocolo:** Envolveu a RSL e o desenvolvimento do PCI.
- **Fase II - Aquisição e Preparação de Dados:** Incluiu a coleta do BD abrangente e a anotação das classes morfológicas.
- **Fase III - Experimentação e Avaliação:** Consistiu na implementação e comparação de múltiplas arquiteturas de DL para estabelecer o *benchmark*.

## 1.5 ESTRUTURA DA DISSERTAÇÃO

O presente trabalho está organizado em uma sequência lógica de seis capítulos, além dos elementos pós-textuais (Referências e Apêndices), conforme descrito a seguir:

- **Capítulo 1 (Introdução):** Apresenta o contexto da pesquisa, a motivação e justificativa do trabalho, a hipótese central, os objetivos, a abordagem metodológica adotada e, por fim, a estrutura da Dissertação.
- **Capítulo 2 (Fundamentação Teórica):** Discute os conceitos teóricos essenciais que sustentam a pesquisa.
- **Capítulo 3 (Trabalhos Correlatos):** Apresenta a RSL de forma concisa.
- **Capítulo 4 (Desenvolvimento):** Detalha o pipeline metodológico da pesquisa.
- **Capítulo 5 (Resultados):** Apresenta os resultados da experimentação com as múltiplas arquiteturas de DL.
- **Capítulo 6 (Considerações Finais):** Sumariza as principais conclusões do trabalho, valida a hipótese central com base nos resultados, discute as limitações da pesquisa e propõe as diretrizes para trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

O presente capítulo visa delinear os fundamentos teóricos e conceitos essenciais que dão suporte à pesquisa, estabelecendo o referencial necessário para a compreensão integral da metodologia e da análise de resultados.

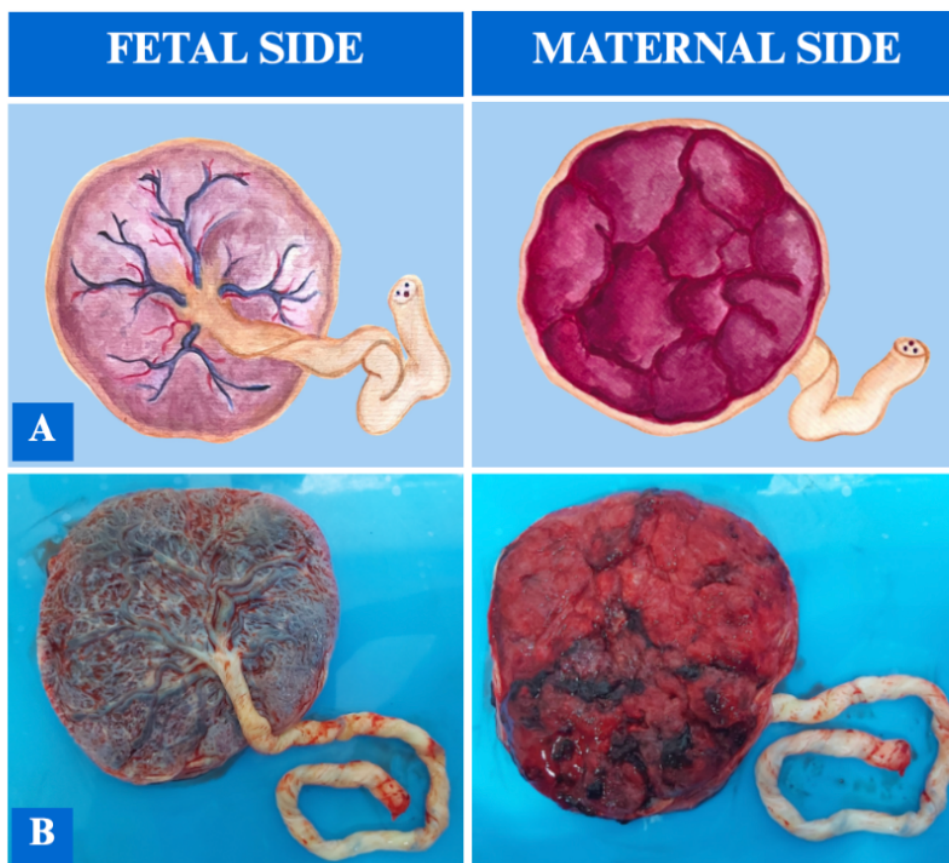
### 2.1 PLACENTA

A placenta é um órgão complexo e transitório, essencial para a manutenção da gestação e para a saúde materno-fetal, atuando como interface vital para a troca gasosa, nutrientes, hormônios e resíduos entre a circulação materna e fetal. Sua análise pós-parto é crucial, pois alterações morfológicas ou patológicas estão frequentemente associadas a desfechos adversos da gravidez, como restrição de crescimento fetal, pré-eclâmpsia e óbito fetal (ROBERTS, 2008; PTACEK *et al.*, 2014).

Para uma triagem macroscópica eficaz, é fundamental compreender suas principais estruturas visíveis. Tipicamente, a placenta é um órgão discóide composto por três componentes principais:

- **Disco Placentário:** Constitui a massa principal do órgão, comumente medindo entre 15 a 20 cm de diâmetro. Apresenta duas faces distintas:
  - **Face Fetal (Corial):** Caracteriza-se por ser lisa e brilhante, coberta pela membrana amniótica. É por esta face que os vasos coriais (artérias e veias) se irradiam do cordão umbilical até a periferia, formando a "árvore"vascular.
  - **Face Materna (Basal):** Possui uma aparência lobulada e irregular, sendo dividida em cotilédones. Esta é a face que estava em contato direto com o útero e é inspecionada quanto à sua integridade e à presença de lesões.
- **Cordão Umbilical:** Representa a conexão vital entre o feto e a placenta, fundamental para o transporte sanguíneo. A avaliação da sua inserção no disco (central, marginal ou velamentosa) é um critério fundamental de triagem. A inserção velamentosa, por exemplo, é um achado que frequentemente exige exame anatomopatológico aprofundado devido aos riscos associados (KHONG *et al.*, 2016).
- **Membranas:** O âmnio e o córion são as membranas que circundam o feto. Devem ser cuidadosamente inspecionadas quanto à sua integridade e ao ponto de ruptura, que podem fornecer informações sobre o parto e possíveis infecções.

Figura 4 – Análise macroscópica da placenta através de representações esquemáticas (A) e imagens reais (B). Lado fetal brilhante com vasos sanguíneos dispostos em padrão arboriforme e cordão umbilical conectado ao centro do disco. Lado materno intacto e discoide, com textura esponjosa cor de vinho coberta por tecido uterino (decídua) e contendo alguns coágulos.



Fonte: Elaborado pelos autores (SILVA, B. Q. T. *et al.*, 2025).

### 2.1.1 Análise Macroscópica e Alterações Placentárias

A avaliação macroscópica da placenta é o exame visual do órgão realizado a olho nu imediatamente após o parto. Este procedimento permite a identificação de alterações estruturais no disco placentário e no cordão umbilical, bem como anomalias na coloração e na integridade das membranas amnióticas. Tais observações são cruciais para a investigação de eventos patológicos ocorridos durante a gestação e são frequentemente determinantes para a decisão de solicitar um exame anatomopatológico aprofundado (KHONG *et al.*, 2016; SILVA, T. *et al.*, 2016).

No entanto, apesar de sua importância, a análise macroscópica tradicional enfrenta desafios significativos que impactam sua eficácia e consistência. A subjetividade inerente à avaliação manual realizada por diferentes profissionais de saúde e a variação na adoção de protocolos hospitalares podem levar a inconsistências na triagem e à subdetecção de patologias placentárias. Conforme destacado no Capítulo 1, essa falta de padronização compromete a detecção precoce e, conseqüentemente, a tomada

de decisões clínicas adequadas (SILVA, T. *et al.*, 2016). É neste cenário de limitação que a necessidade de ferramentas de triagem eficientes, rápidas e de baixo custo se torna evidente, pavimentando o caminho para o desenvolvimento de metodologias automatizadas.

### 2.1.2 Exame Anatomopatológico

O exame anatomopatológico da placenta representa o padrão ouro para o diagnóstico definitivo de patologias placentárias, sendo realizado por patologistas através da análise microscópica dos tecidos. Este exame detalhado é indispensável para elucidar a causa de complicações obstétricas e neonatais, como restrição de crescimento fetal, pré-eclâmpsia e óbito fetal, além de fornecer informações cruciais para o aconselhamento genético e o manejo de futuras gestações (KHONG *et al.*, 2016; ROBERTS, 2008).

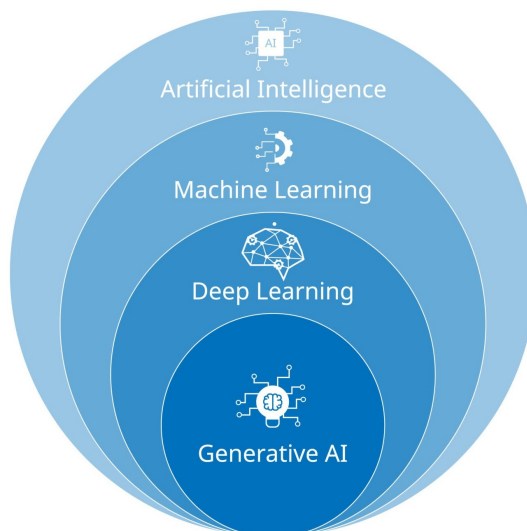
Contudo, a solicitação sistemática deste exame para todas as placentas é inviável na prática clínica, principalmente devido ao seu alto custo e à significativa demanda operacional que impõe aos laboratórios de patologia. Assim, a decisão de encaminhar uma placenta para análise laboratorial é frequentemente baseada em critérios clínicos e nas características morfológicas observadas na avaliação macroscópica inicial (SILVA, T. *et al.*, 2016).

## 2.2 INTELIGÊNCIA ARTIFICIAL (IA)

A IA pode ser definida como o esforço contínuo de conferir conhecimento humano a sistemas computacionais. O propósito central da IA é capacitar máquinas a pensar, aprender e agir de maneira autônoma, simulando a inteligência natural para executar tarefas complexas e resolver problemas. A vasta maioria dos sistemas de IA concentra-se em emular a cognição humana para abordar desafios de alta complexidade (JAKHAR; KAUR, 2020).

Atualmente, dentre as diversas abordagens para desenvolver essa capacidade de resolução inteligente de problemas em computadores, o Machine Learning (ML) e o DL representam as técnicas mais proeminentes. A Figura 5 ilustra claramente a hierarquia e a relação entre esses campos, apresentando a IA como o domínio mais amplo, que engloba o ML, e este, por sua vez, engloba o DL como uma subárea especializada. Sua aplicação é transversal, abrangendo virtualmente todos os campos da ciência. No contexto da medicina, em particular, as metodologias de IA são intensamente utilizadas, desempenhando um papel crucial tanto na automação de processos quanto no auxílio ao diagnóstico em exames de imagens médicas.

Figura 5 – Hierarquia e Relação entre os Principais Campos da IA.



Fonte: Elaborado pela autora (2025).

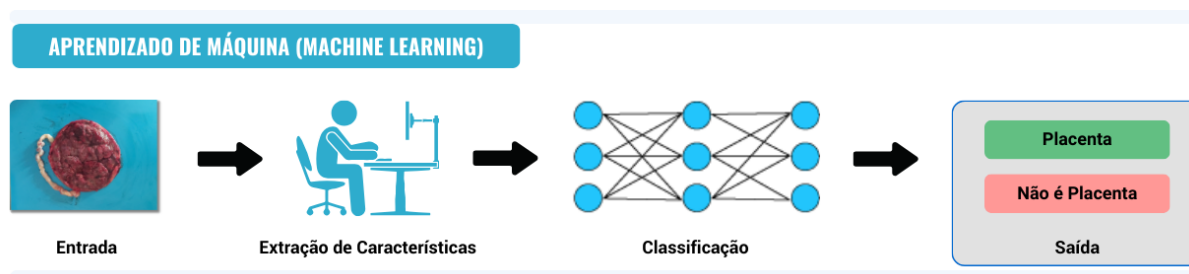
### 2.2.1 Machine Learning (ML)

O ML é um campo da ciência da computação focado no desenvolvimento de modelos preditivos. Para tal, o ML emprega algoritmos e utiliza análises estatísticas e computacionais para extrair padrões de dados e solucionar problemas práticos (JAKHAR; KAUR, 2020).

Em sua essência, o ML representa o desenvolvimento de algoritmos capazes de aprender diretamente a partir dos dados, sem a necessidade de seguir regras de programação estritamente definidas e fixas (SAMUEL, 2000).

Esta técnica é amplamente utilizada, notavelmente em tarefas de classificação, onde o sistema aprende a categorizar novas informações com base em exemplos anteriores.

Figura 6 – Fluxo base ML aplicado a classificação de imagem.



Fonte: Elaborado pela autora (2025).

### 2.2.2 Deep Learning (DL)

O DL é uma subárea do ML que se destaca pelo uso de Redes Neurais Artificiais (RNA). Essas redes são concebidas para imitar a estrutura e o funcionamento do cérebro humano, sendo formadas por múltiplas camadas interconectadas.

No DL, a rede recebe informações e, com base no conhecimento que adquire através de vastos conjuntos de exemplos, é capaz de gerar uma saída (JAKHAR; KAUR, 2020; ZHANG, W. J. *et al.*, 2018). Dentro dessa estrutura:

- Cada neurônio artificial resolve uma pequena parte do problema.
- O conhecimento é transmitido entre os neurônios interconectados, permitindo que a rede, em seu conjunto, resolva o problema inicial e forneça uma resposta de saída.

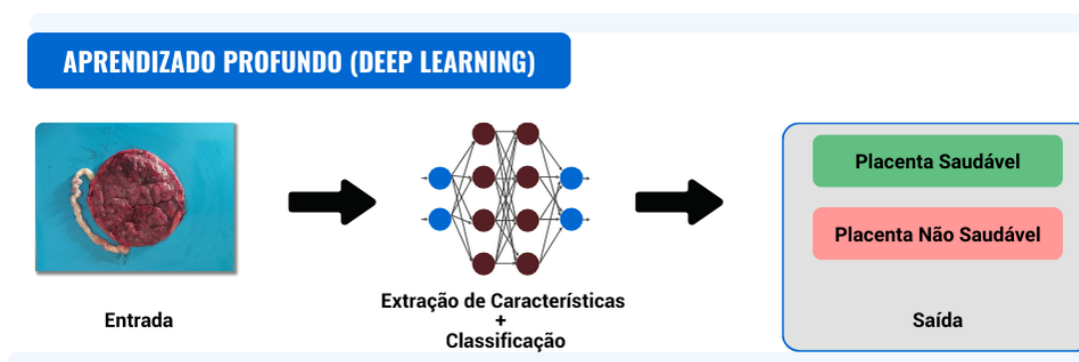
#### Diferenciais do Deep Learning:

Os algoritmos de DL apresentam uma distinção crucial em relação aos algoritmos tradicionais de ML:

- **Aprendizado de Características (Feature Learning):** O DL possui a capacidade intrínseca de aprender e extrair características relevantes diretamente dos dados brutos (aprendizado de maneira intuitiva), mesmo que esses dados não sejam estruturados.
- **Independência de Pré-definição:** O DL não requer que as características do conjunto de dados sejam previamente inseridas ou definidas por humanos. A rede é capaz de descobrir hierarquicamente as representações mais importantes para a tarefa, como ilustrado em uma classificação de imagem .

Atualmente, o DL é o campo de maior proeminência na área de VC e é a técnica dominante na análise de imagens médicas.

Figura 7 – Fluxo base DL aplicado a classificação de imagem.

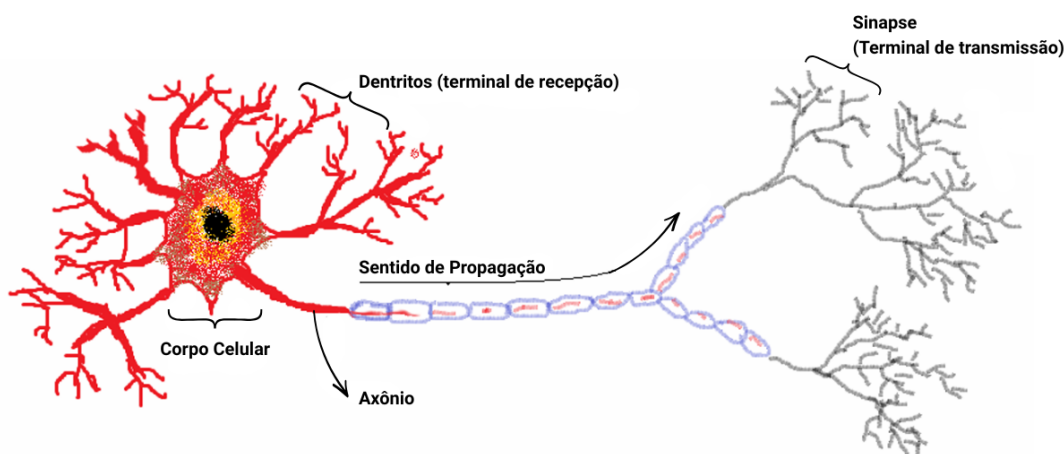


Fonte: Elaborado pela autora (2025).

### 2.2.3 Redes Neurais Artificiais (RNA)

As RNA consistem em um modelo computacional cuja inspiração reside na estrutura e no funcionamento biológico do cérebro humano.

Figura 8 – Representação do neurônio biológico.



Fonte: Elaborado pela autora (2025).

A RNA é composta por um conjunto de unidades de processamento interconectadas, frequentemente denominadas neurônios ou nós, que se organizam em camadas sequenciais:

- **Camada de Entrada (Input Layer):** Responsável por receber os dados brutos de entrada do problema.
- **Camada(s) Oculta(s) (Hidden Layer(s)):** Onde ocorre o processamento central, focado na extração e codificação de padrões nos dados.
- **Camada de Saída (Output Layer):** Produz o resultado final da tarefa, como uma predição de classe ou valor.

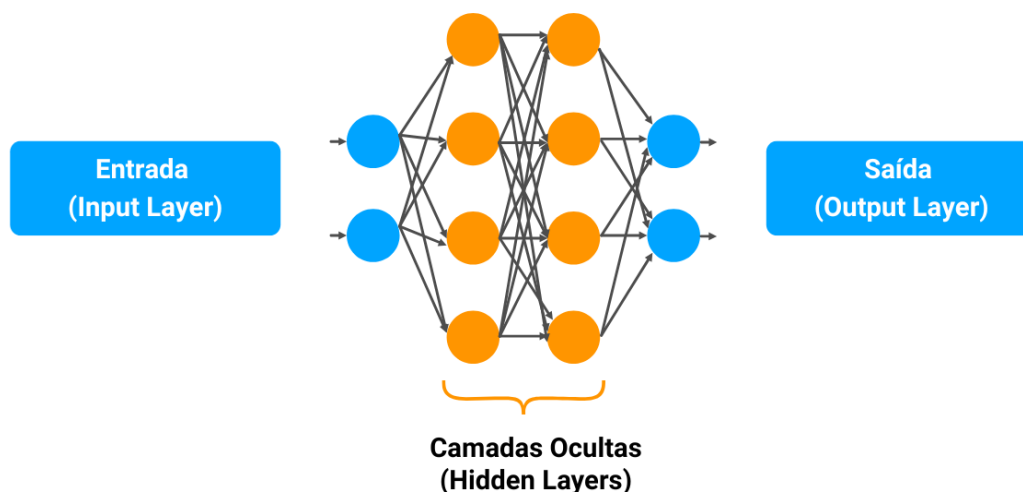
A transmissão de informação entre neurônios adjacentes é modulada por um peso associado a cada conexão e um bias, permitindo o ajuste fino do sinal de saída. O processo de treinamento da rede é a otimização iterativa desses pesos e biases, visando adequar o modelo ao conjunto de entradas e saídas observadas.

#### Limitação em Processamento de Imagens:

Um desafio inerente às arquiteturas de RNA convencionais reside na sua estrutura frequentemente totalmente conectada (fully connected), onde cada nó se liga a todos os nós da camada subsequente. Ao processar dados visuais, isso obriga a transformar a matriz 2D de pixels em um vetor unidimensional.

Essa vetorização acarreta duas grandes desvantagens (LECUN *et al.*, 1998):

Figura 9 – Representação RNA.



Fonte: Elaborado pela autora (2025).

- **Perda de Informação Espacial:** A relação crucial entre pixels vizinhos é desconsiderada, pois eles são tratados da mesma forma que pixels distantes.
- **Explosão de Parâmetros:** O número de pesos na rede aumenta drasticamente, elevando o custo computacional e dificultando o treinamento.

Para superar essas deficiências no tratamento de dados com estrutura espacial, foram desenvolvidas as CNN.

#### 2.2.4 Redes Neurais Convolucionais (CNN)

As CNN representam uma classe de redes neurais otimizadas e especializadas no processamento de dados que possuem uma estrutura espacial, sendo o exemplo mais notável as imagens. Sua criação teve como objetivo primordial mitigar as limitações das RNA tradicionais ao preservar as relações espaciais entre os pixels e, simultaneamente, reduzir o volume de parâmetros do modelo (LECUN *et al.*, 1998). As RNA convencionais, com suas camadas totalmente conectadas, tendem a perder a informação contextual e espacial inerente às imagens, transformando-as em vetores unidimensionais. Além disso, essa abordagem resulta em uma "explosão de parâmetros", tornando o treinamento computacionalmente proibitivo para imagens de alta resolução.

A inspiração biológica para a arquitetura das CNN remonta aos estudos de Hubel e Wiesel em 1962, que demonstraram como neurônios no córtex visual de seres vivos respondem a estímulos visuais específicos, como bordas e linhas em determinadas orientações. Essa descoberta impulsionou a ideia de que o reconhecimento visual hierárquico poderia ser emulado computacionalmente. As CNN mimetizam essa orga-

nização do córtex visual, com neurônios artificiais que respondem a padrões locais da imagem e constroem representações progressivamente mais complexas.

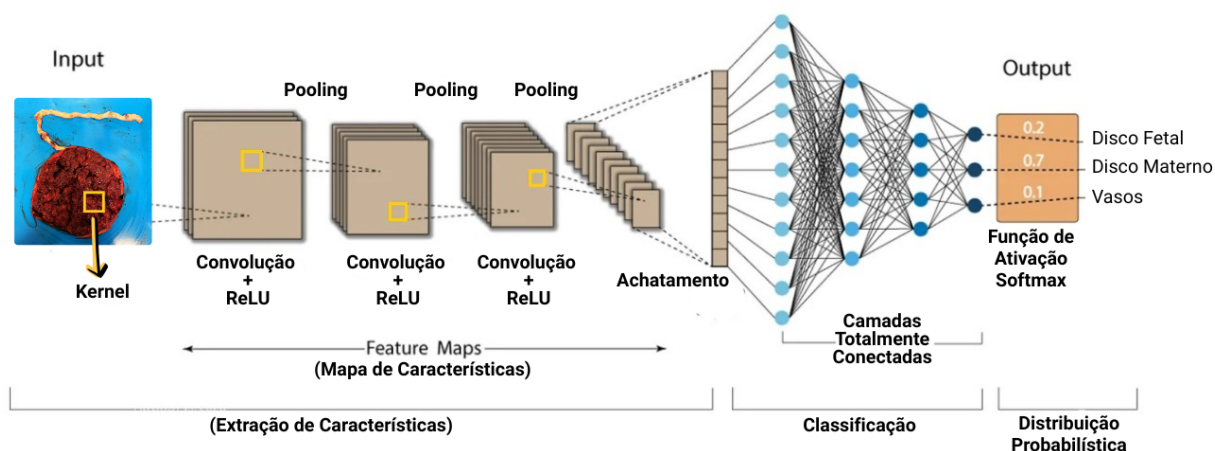
Uma CNN para classificação ou reconhecimento de imagem recebe uma entrada que é tipicamente uma matriz tridimensional (altura, largura e profundidade, sendo esta última determinada pela quantidade de canais de cores, como RGB) (Lecun *et al.*, 1998). Ao longo de sua arquitetura, a rede utiliza filtros (também chamados de *kernels* ou detectores de características) que realizam operações de convolução. Estes filtros percorrem localmente pequenas regiões da imagem de entrada, identificando padrões hierárquicos que evoluem de detalhes simples (como bordas e texturas) a formas e características complexas. O principal benefício é a capacidade de aprender esses filtros automaticamente durante o treinamento, adaptando-os para as características mais relevantes do conjunto de dados.

A arquitetura de uma CNN é tipicamente composta por uma sequência de camadas, sendo as mais comuns:

- **Camadas Convolucionais:** São o núcleo das CNN. Cada neurônio nesta camada se conecta apenas a uma pequena região da entrada da camada anterior. Essa operação de convolução aplica um filtro (*kernel*) sobre a imagem de entrada, gerando um mapa de características (*feature map*) que destaca a presença de padrões específicos, como bordas, texturas ou formas. Múltiplos filtros podem ser aplicados para extrair diferentes tipos de características.
- **Camadas de Agregação (*Pooling* ou *Subamostragem*):** Usualmente inseridas entre camadas convolucionais, as camadas de agregação têm como objetivo reduzir a dimensionalidade espacial dos mapas de características, diminuir a quantidade de parâmetros e computação, e tornar o modelo mais robusto a pequenas variações e deslocamentos na imagem. As operações mais comuns são o *max pooling* e o *average pooling*, que selecionam o valor máximo ou a média, respectivamente, dentro de uma janela deslizante.
- **Camadas Totalmente Conectadas (*Fully Connected Layers*):** Após várias camadas convolucionais e de agregação extraírem e refinarem as características da imagem, os mapas de características resultantes são "achatados" em um vetor unidimensional. Este vetor é então passado para uma ou mais camadas totalmente conectadas, que funcionam como uma RNA tradicional. Estas camadas são responsáveis por realizar a classificação final ou detecção, utilizando as características de alto nível aprendidas pelas camadas anteriores.

A combinação dessas camadas permite que as CNN aprendam representações complexas e abstratas diretamente dos dados brutos, tornando-as extremamente eficazes em tarefas como classificação, detecção e segmentação de imagens.

Figura 10 – Representação CNN.



Fonte: Elaborado pela autora (2025).

### 2.3 VISÃO COMPUTACIONAL (VC)

A VC é um domínio da ciência que engloba o conjunto de métodos e técnicas que visam capacitar sistemas computacionais a processar, analisar e interpretar informações provenientes de imagens e vídeos.

A VC pode ser categorizada em duas abordagens principais (WANGENHEIM, Aldo von; COMUNELLO; RICHA, 2013):

- **Visão Clássica:** Baseada em métodos numéricos, operações de morfologia matemática e processamento de imagem tradicional.
- **VC Baseada em IA:** Utiliza as técnicas de ML e DL para a interpretação e inferência de dados visuais.

As principais aplicações e tarefas da VC incluem:

- **Classificação de Imagens:** Determina a categoria ou o rótulo da imagem em sua totalidade (Ex: "Esta é uma imagem de paisagem?").
- **Detecção de Objetos:** Localiza objetos de interesse na imagem e os delimita com uma caixa delimitadora (bounding box), atribuindo-lhes um rótulo de classe.
- **Segmentação Semântica:** O objetivo é classificar cada pixel individual da imagem em uma classe pré-definida. Múltiplos objetos pertencentes à mesma classe são tratados como uma única entidade segmentada.
- **Segmentação por Instância:** Combina a detecção de objetos com a segmentação semântica, permitindo diferenciar e individualizar múltiplos objetos

da mesma classe (e.g., separar o pixel exato de um tumor de outro, se houver dois).

A área de VC, impulsionada pela aplicação de CNN — um subtipo de DL —, é a responsável pelos maiores avanços na análise computacional de imagens médicas. As CNN reduziram drasticamente a necessidade de engenharia de features manual e o número de parâmetros, acelerando os estudos em diversas modalidades, como raio-x, ressonância magnética, microscopia e dermatologia (PURI *et al.*, 2020).

### 3 TRABALHOS CORRELATOS

Para a identificação e análise de trabalhos relacionados, foi conduzida uma RSL seguindo os protocolos estabelecidos por Kitchenham (KITCHENHAM, 2004), aplicada às principais bases de dados científicas que suportam operadores booleanos na construção de strings de busca.

O detalhamento integral da metodologia da RSL (que engloba o objetivo, a estratégia e os termos de busca, o processo de seleção com seus critérios de inclusão e exclusão, o detalhamento e a análise crítica dos artigos primários, a discussão dos resultados e a conclusão) encontra-se documentado no **Apêndice A** deste trabalho, no formato de Relatório Técnico (LOPES; SILVA, B. Q. T.; WANGENHEIM, A. von, 2025b).

#### 3.1 DEFINIÇÃO E EXECUÇÃO DA BUSCA

A estratégia de busca foi estruturada para mapear o estado da arte em métodos automatizados de análise macroscópica placentária utilizando imagens fotográficas, com ênfase em técnicas de VC e DL. Foram consultadas cinco bases de dados científicas: IEEE Xplore, ScienceDirect, ACM Digital Library, PubMed e Scielo, no período de 2019 a 2024, utilizando uma estratégia de busca consolidada, que incluiu termos-chave como: "*placenta*", "*pathology*", "*deep learning*", "*segmentation*", "*photo image analysis*".

Tabela 1 – Resumo da definição do protocolo utilizado na busca.

|  |
|--|
| <b>Base de Dados</b>   |
| IEEE Xplore<br>ScienceDirect<br>ACM Digital Library<br>PubMed<br>Scielo  |
| <b>Período</b>   |
| 2019 a 2024  |
| <b>String Geral de Busca</b>   |
| <i>(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).</i> |

Fonte: Elaborado pela autora (2025).

A string de busca foi modelada para capturar trabalhos que aplicassem técnicas de DL na análise de imagens placentárias macroscópicas, excluindo estudos baseados exclusivamente em exames clínicos de imagem (ressonância magnética, ultrassonografia) sem componente fotográfico. O intervalo temporal restrito (2019-2024)

justifica-se pela evolução recente e acelerada das arquiteturas de CNN aplicadas ao domínio médico.

A execução do protocolo resultou inicialmente em 459 artigos. Após análise de título e resumo, foram selecionados 39 trabalhos potencialmente relevantes para leitura completa. Aplicando-se rigorosamente os critérios de inclusão e exclusão durante a análise integral dos textos, foram identificados cinco estudos diretamente alinhados ao escopo desta pesquisa.

Tabela 2 – Resultados e Processo de Filtragem da Busca.

| Etapa de Seleção               | Número de Artigos |
|--------------------------------|-------------------|
| Identificação Inicial          | 459               |
| Triagem (Título + Resumo)      | 39                |
| Elegibilidade (Texto Completo) | 5                 |

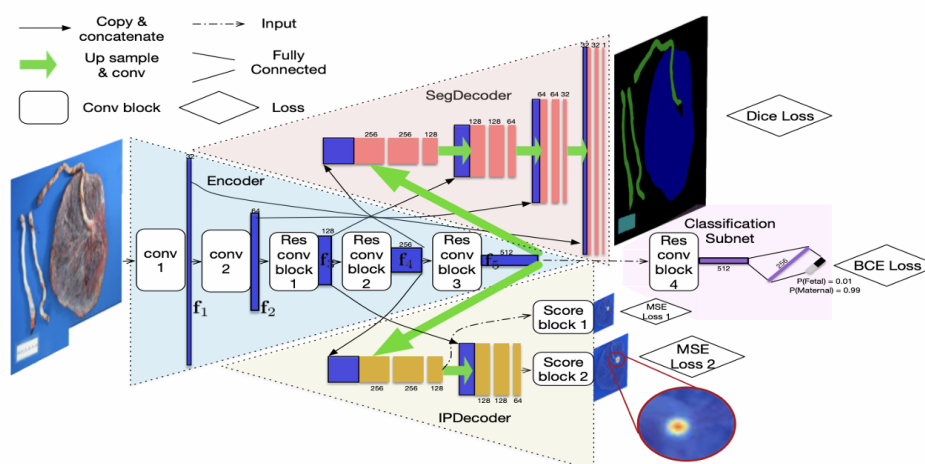
Fonte: Elaborado pela autora (2025).

### 3.2 SÍNTESE DOS RESULTADOS

Os cinco trabalhos selecionados representam o conjunto de pesquisas que abordam especificamente a análise automatizada de placentas através de imagens fotográficas utilizando DL. A análise comparativa destes estudos revela:

**PlacentaNet: Automatic Morphological Characterization of Placenta Photos with Deep Learning:** Primeiro framework dedicado à segmentação e classificação automatizada de estruturas placentárias, utilizando arquitetura baseada em U-Net modificada para identificação de regiões anatômicas e detecção de alterações morfológicas (CHEN *et al.*, 2019).

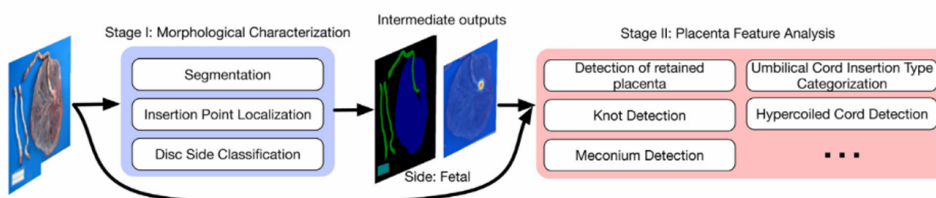
Figura 11 – Arquitetura do PlacentaNet.



Fonte: CHEN; WU et al. (2019).

**AI-PLAX: AI-based placental assessment and examination using photos:** Sistema integrado para análise placentária que combina segmentação semântica, localização de cordão umbilical e classificação de anomalias, demonstrando viabilidade clínica em ambiente hospitalar (CHEN *et al.*, 2020).

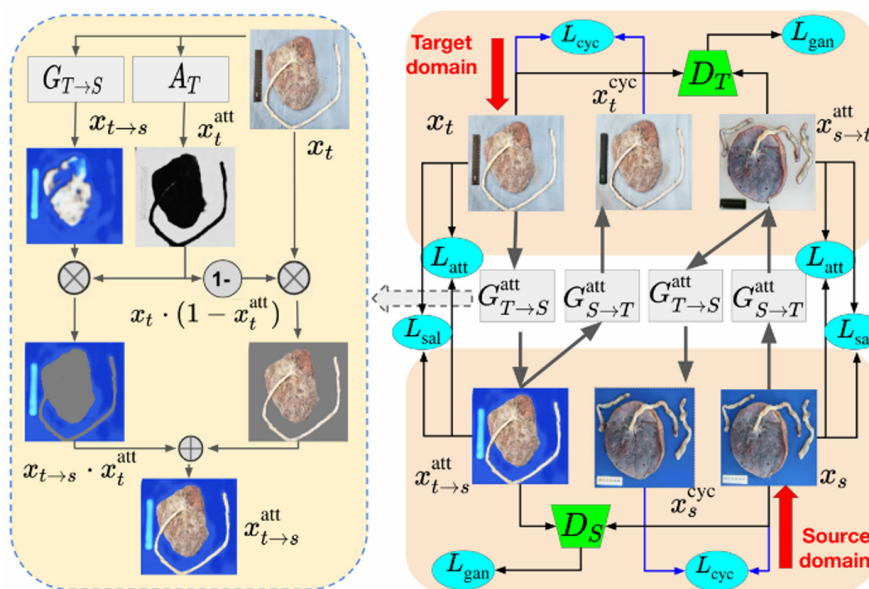
Figura 12 – Diagrama esquemático do framework proposto de dois estágios para avaliação e exame automatizados da placenta utilizando fotos.



Fonte: CHEN; WU *et al.* (2020).

**Multi-region saliency-aware learning for cross-domain placenta image segmentation:** Abordagem que incorpora mecanismos de atenção para priorizar regiões de interesse clínico, melhorando a interpretabilidade dos modelos e reduzindo falsos positivos (ZHANG, Z. *et al.*, 2020).

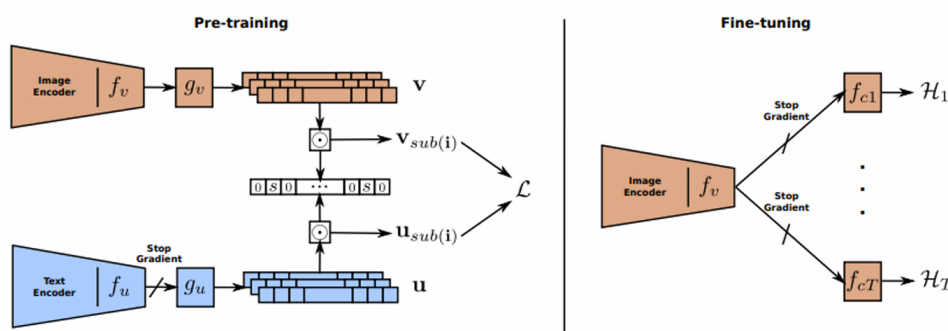
Figura 13 – Pipeline MSL (Aprendizado de Saliência Multi-regional) proposto.



Fonte: ZHANG *et al.*, (2020).

**Vision-Language Contrastive Learning Approach to Robust Automatic Placenta Analysis Using Photographic Images:** Método inovador que integra informações visuais e textuais (laudos clínicos) através de aprendizado contrastivo, permitindo melhor alinhamento entre características visuais e terminologia médica (PAN *et al.*, 2022).

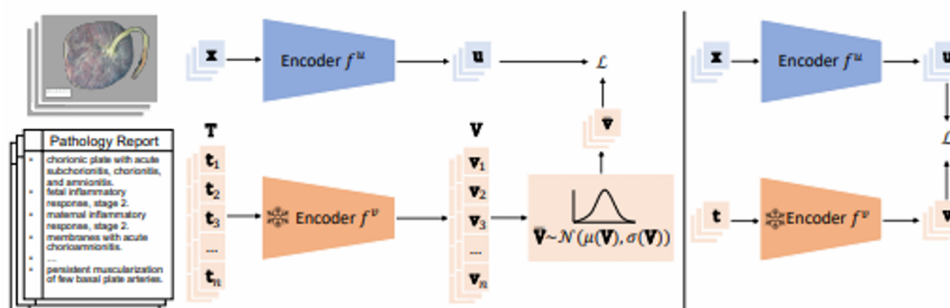
Figura 14 – Pipeline VLC (Aprendizado Contrastivo de Visão e Linguagem) proposto.



Fonte: PAN et al., (2022).

**Enhancing Automatic Placenta Analysis through Distributional Feature Recomposition in Vision-Language Contrastive Learning:** Trabalho mais recente que propõe técnicas de adaptação de domínio para melhorar generalização entre diferentes protocolos de captura de imagem e populações clínicas (PAN *et al.*, 2023).

Figura 15 – Método avançada de VLC proposto para analisar automaticamente imagens fotográficas de placentas.



Fonte: PAN et al., (2023).

### 3.3 DISCUSSÃO E POSICIONAMENTO DA PESQUISA

A análise dos cinco trabalhos resultantes (AI-PLAX, 2020; PlacentaNet, 2019; Multi-region Saliency-Aware Learning, 2020; Vision-Language Contrastive Learning, 2022; Enhancing Automatic Placenta Analysis, 2023) revelou uma evolução consistente na aplicação de técnicas de Deep Learning à Placentologia.

Contudo, a RSL evidencia as seguintes lacunas críticas que limitam a adoção clínica imediata das soluções propostas:

- **Ausência de sistemas de triagem automatizada:** Os trabalhos identificados concentram-se predominantemente em diagnóstico automatizado de alterações específicas, não abordando o problema de triagem inicial para solicitação de exame anatomopatológico - foco central desta dissertação.

- **Qualidade dos Dados:** Os modelos existentes foram treinados majoritariamente em imagens de alta qualidade, obtidas em ambientes controlados (laboratório de patologia), com iluminação e enquadramento padronizados.
- **Ambiente Real:** Existe uma escassez de trabalhos que utilizem imagens capturadas em condições clínicas adversas (como a sala de parto), onde a iluminação inconsistente, a presença de artefatos e a não-padronização do enquadramento são desafios reais.
- **Limitação de datasets públicos:** Todos os estudos utilizam conjuntos de dados proprietários, dificultando reprodutibilidade e comparação direta de resultados, além de limitar a validação externa dos modelos propostos.
- **Validação clínica restrita:** A maioria dos trabalhos apresenta validação técnica robusta (métricas computacionais), porém com validação clínica limitada quanto à aplicabilidade em fluxos de trabalho hospitalares reais e impacto na tomada de decisão médica.

Estas lacunas justificam a relevância e a originalidade desta pesquisa, que propõe uma abordagem de triagem automatizada para a identificação de alterações macroscópicas que demandem análise especializada. A solução é duplamente inovadora: fundamentada em um PCI padronizado para garantir a robustez dos dados e posicionada como uma ferramenta de suporte à decisão clínica em contexto de maternidades e centros obstétricos.

## 4 DESENVOLVIMENTO

O presente capítulo detalha a metodologia e o pipeline de desenvolvimento implementado para atingir os objetivos propostos. O pipeline de desenvolvimento está estruturado em três eixos principais, apresentados sequencialmente: o processo de aquisição e preparação de dados (Seção 4.1), o ambiente e a configuração da experimentação (Seção 4.2) e a formalização das métricas de avaliação (Seção 4.3). O rigor na descrição desses materiais e métodos é essencial para garantir a reprodutibilidade técnica e a validade dos resultados apresentados no Capítulo 5.

### 4.1 MATERIAIS E AQUISIÇÃO DE DADOS

#### 4.1.1 Protocolo de Captura de Imagens (PCI)

O principal desafio metodológico para a triagem placentária automatizada reside na variabilidade e na não-padronização das imagens capturadas no ambiente clínico (sala de parto), conforme evidenciado no Capítulo 3. Essa inconsistência nos dados compromete a robustez e a capacidade de generalização dos modelos de DL.

Para mitigar essa limitação, foi desenvolvido um PCI de Placentas na Sala de Parto, documentado no Relatório Técnico (LOPES; SILVA, B. Q. T.; WANGENHEIM, A. von, 2025a). Ressalta-se que todo o processo de aquisição de imagens respeitou rigorosamente as diretrizes éticas e foi aprovado pelo Comitê de Ética em Pesquisa (CEP/CONEP), sob o número CAAE 86362524.7.0000.0121.

O objetivo central desse protocolo é propor uma metodologia padrão a ser executada pela equipe assistente, utilizando dispositivos móveis, que garanta a obtenção de imagens com um mínimo de informações necessárias para o reconhecimento e a classificação automatizada de alterações macroscópicas. A padronização de elementos como contraste de fundo, enquadramento e iluminação é essencial para a eficácia do treinamento das redes neurais.

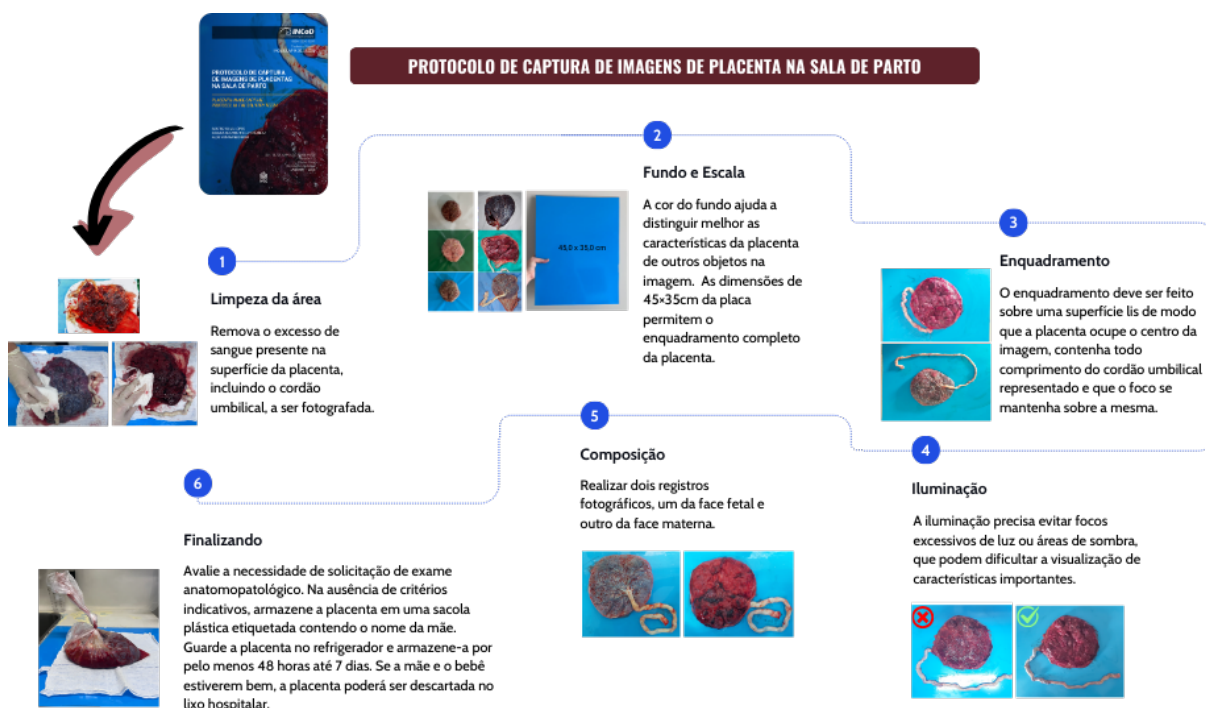
O processo metodológico do PCI está estruturado em uma sequência de seis etapas principais, conforme ilustrado na Figura 16.

##### 4.1.1.1 Detalhamento de Execução do PCI

A execução do PCI focou em transformar o ambiente de coleta em um cenário controlável, garantindo a qualidade do dado como entrada para os modelos de DL.

O preparo da cena teve início com a completa remoção da placenta do útero materno e a finalização dos cuidados neonatais imediatos, garantindo um ambiente calmo e seguro para a manipulação da peça. A placenta foi posicionada sobre uma superfície plana e limpa, preferencialmente uma placa de acrílico ( $45,0 \times 35,0 \times 0,5$  cm) recoberta por adesivo azul vinil. Essa superfície proporciona o contraste visual ade-

Figura 16 – Fluxo de Execução do PCI.



Fonte: Elaborado pela autora (2025).

quado, mitigando a interferência de fundo e facilitando a padronização entre amostras. Na ausência desta placa, campos cirúrgicos de coloração neutra foram utilizados como alternativa viável.

Antes da fotografia, o excesso de sangue foi cuidadosamente removido tanto da superfície da placenta quanto do cordão umbilical, com uso de gazes ou compressas estéreis. Cada espécime foi fotografado em duas faces anatômicas distintas — a face fetal (contendo os vasos arboriformes, membranas e inserção do cordão umbilical) e a face materna (com os cotilédones e septações), expondo-se o disco placentário mediante retração manual das membranas amnióticas.

O enquadramento foi realizado com o dispositivo fotográfico posicionado perpendicularmente ( $90^\circ$ ) em relação à superfície. Essa angulação garante a captura integral da placenta com seus principais componentes estruturais, minimizando a distorção espacial. As dimensões da placa acrílica também funcionaram como referência interna de escala.

As condições de iluminação foram ajustadas para evitar sombras, brilho excessivo ou reflexos. Sempre que necessário, as luzes cirúrgicas de teto foram desligadas, sendo utilizada iluminação ambiente ou difusa para garantir a homogeneidade visual. As fotografias foram capturadas com smartphones (Android ou iOS), utilizando câmeras traseiras com resolução mínima de 12 megapixels.

Evitaram-se imagens com interferências visuais (p. ex., mãos com luvas, ins-

trumentais ou manchas excessivas). Todos os registros passaram por revisão técnica, considerando critérios de clareza, visibilidade anatômica e conformidade com o protocolo. O contraste obtido entre as abordagens padronizada e não padronizada é evidenciado na Figura 17.

Figura 17 – Exemplo de Imagem de Placenta Padronizada (segundo o PCI) e Imagem Não Padronizada (sem a aplicação do PCI).



Fonte: Elaborado pela autora (2025).

#### 4.1.2 Conjunto de Imagens (Dataset)

O Conjunto de Imagens (Dataset) constitui o material empírico base para a experimentação. O presente estudo utilizou um conjunto de dados composto por 156 imagens originais capturadas imediatamente após o parto em centros obstétricos e maternidades parceiras. As imagens foram coletadas em dois contextos distintos: uma parte seguiu condições hospitalares rotineiras, e a outra foi obtida conforme o PCI. Parte dessas imagens foi coletada no serviço de Obstetrícia do Hospital Universitário Polydoro Ernani de São Thiago (HU-UFSC/EBSERH), seguindo o protocolo padronizado (LOPES; SILVA, B. Q. T.; WANGENHEIM, A. von, 2025a), enquanto outras foram adquiridas em diferentes instituições.

##### 4.1.2.1 Processamento e Classes

As anotações focaram nas principais estruturas morfológicas clinicamente relevantes da placenta. No entanto, para os experimentos, apenas as classes com quantidade suficiente de exemplos visuais foram utilizadas. O Quadro 1 apresenta a distribuição das anotações por classe morfológica.

As imagens foram organizadas e anotadas na plataforma Roboflow, que oferece recursos interativos para gerenciamento de conjuntos e exportação.

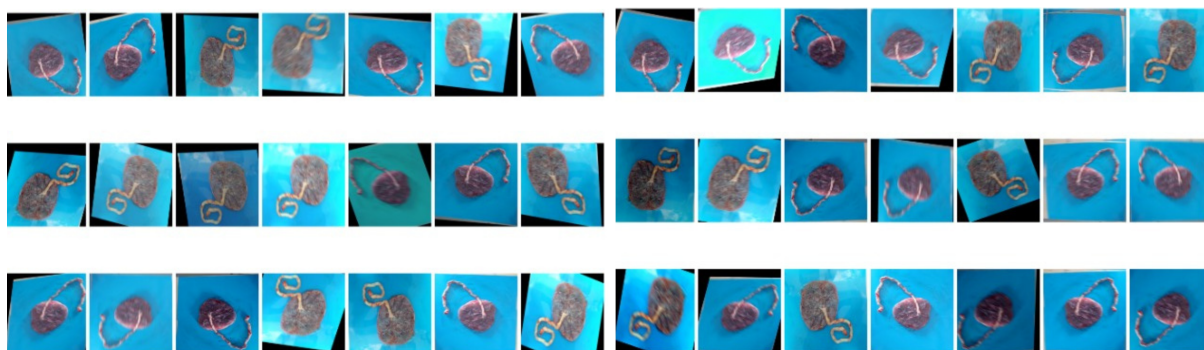
Quadro 1 – Distribuição da quantidade de instâncias rotuladas por classe morfológica utilizada no treinamento.

| Classe Morfológica       | Instâncias Anotadas |
|--------------------------|---------------------|
| cordão_umbilical_fetal   | 83                  |
| cordão_umbilical_materno | 72                  |
| disco_discoide_fetal     | 56                  |
| disco_discoide_materno   | 29                  |
| disco_irregular_fetal    | 7                   |
| disco_irregular_materno  | 8                   |
| disco_ovalada_fetal      | 16                  |
| disco_ovalada_materno    | 41                  |
| vasos_arboriforme        | 515                 |

Fonte: Elaborado pela autora (2025).

O conjunto foi inicialmente expandido de 156 para 468 imagens por meio de técnicas de aumento de dados (*data augmentation*) (BUSLAEV *et al.*, 2020). As transformações aplicadas incluíram rotações ( $\pm 15^\circ$ ), inversões horizontal e vertical (flips), distorção (shear  $\pm 10^\circ$ ), e variações de saturação ( $\pm 25\%$ ) e exposição ( $\pm 10\%$ ).

Figura 18 – Exemplos de imagens sintéticas obtidas por *data augmentation*, simulando variações comuns em capturas com dispositivos móveis.



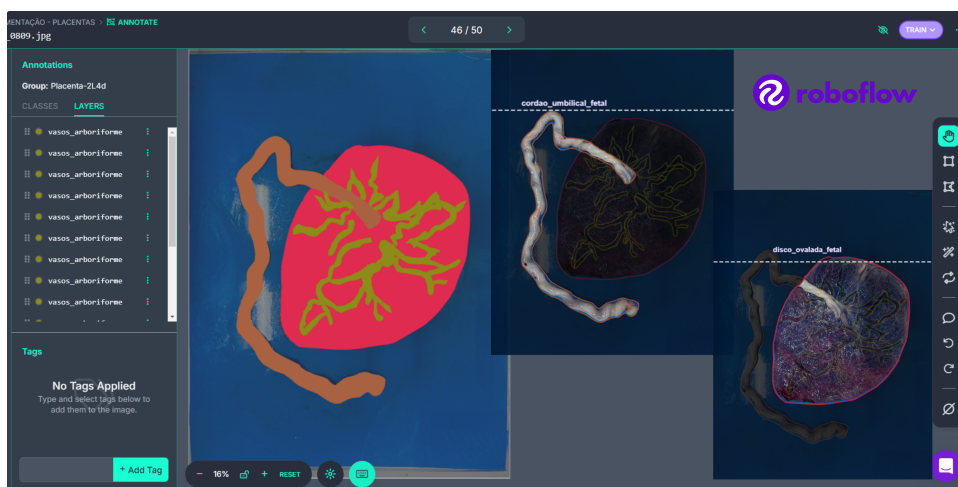
Fonte: Elaborado pela autora (2025).

#### 4.1.3 Processo de Anotação e Exportação das Imagens

O processo de anotação foi realizado manualmente por especialistas (patologistas).

Durante o processo de preparação, identificou-se um desalinhamento em algumas imagens entre os arquivos RGB e suas respectivas máscaras, causado pelos metadados EXIF Orientation. Esse metadado preserva a orientação original da câmera, mas pode gerar inconsistências quando aplicações de VC não interpretam corretamente o campo *orientation*, resultando em imagens e máscaras com eixos trocados (p. ex., pixels armazenados como  $x,y$  exibidos como  $y,x$ ) (DWYER, 2020).

Figura 19 – Exemplo de anotação no Roboflow.



Fonte: Elaborado pela autora (2025).

Para corrigir esse problema e garantir a consistência entre imagem e máscara, aplicou-se o recurso Auto-Orient do Roboflow durante o pré-processamento. Essa função remove os metadados de rotação e reescreve a matriz de pixels na orientação final correta, padronizando todas as imagens independentemente da posição da câmera no momento da captura (DWYER, 2020).

Após essa normalização, as imagens passaram por padronização de resolução para  $1600 \times 1200$  px e conversão para formato PNG sem compressão com perda, garantindo a preservação dos detalhes morfológicos relevantes.

Esse ajuste de Auto-Orient foi aplicado apenas nos experimentos de segmentação semântica (máscaras por classe), pois na tarefa de detecção por instância com a You Only Live Once (YOLO), as anotações consistiam em caixas delimitadoras (bounding boxes) independentes das máscaras (KHANAM; HUSSAIN, 2024), não apresentando o mesmo problema. Os dados foram finalmente exportados nos seguintes formatos para cada tipo de tarefa:

- **Máscaras (MASK):** para os modelos de segmentação semântica, com uma máscara binária por classe;
- **YOLO (TXT):** para a YOLO, seguindo o padrão oficial da arquitetura com coordenadas normalizadas por classe e posição.

## 4.2 CONFIGURAÇÃO DO AMBIENTE

### 4.2.1 Ambiente de Desenvolvimento

Todos os experimentos foram conduzidos em uma estação de trabalho NVIDIA DGX equipada com **GPU H100**, o que possibilitou o treinamento eficiente de redes profundas com imagens em alta resolução. O ambiente de desenvolvimento foi base-

ado em **Python**, utilizando o *framework* PyTorch como base para a construção e o treinamento dos modelos (PASZKE *et al.*, 2019).

A seguir, destacam-se as principais bibliotecas e *frameworks* utilizados, categorizados por sua função:

- **Deep Learning Core:**

- **PyTorch** (PASZKE *et al.*, 2019): *Framework* principal para computação em *tensor* e construção de redes neurais.
- **fastai** (HOWARD; GUGGER, 2020): Utilizada como camada de abstração sobre o PyTorch, empregada na criação e treinamento das U-Nets com *backbones* ResNet e EfficientNet.

- **Pré-processamento e Data Augmentation:**

- **albumentations** (BUSLAEV *et al.*, 2020): Utilizada para *data augmentation* com transformações rápidas e otimizadas para VC.

- **Arquiteturas Específicas e Bibliotecas:**

- **ultralytics** (JOCHER; QIU; CHAURASIA, 2024): Pacote oficial para a implementação e treinamento da arquitetura YOLO.
- **efficientnet-pytorch** (NOGUEIRA, 2021): Biblioteca utilizada para a integração eficiente dos *backbones* EfficientNet nas arquiteturas U-Net.

- **Monitoramento e Análise:**

- **Weights & Biases (wandb)** (BIEWALD, 2020): Ferramenta de rastreamento e monitoramento de métricas de treinamento.
- **lapixdl** (LAPIX, INCOD, UFSC, 2025): Biblioteca nacional utilizada para visualização e avaliação de segmentações médicas.
- **Matplotlib, Pandas e Seaborn** (HUNTER, 2007; MCKINNEY, 2010; WASKOM, 2021): Empregadas na análise exploratória dos dados, geração de gráficos de desempenho e visualização das matrizes de confusão.

#### 4.2.2 Arquiteturas Implementadas

A seleção das arquiteturas de DL foi guiada por um princípio de comparação estratégica, visando equilibrar o desempenho preditivo (precisão técnica) com a viabilidade operacional (custo computacional).

Para isso, esta pesquisa avaliou 14 arquiteturas aplicadas à segmentação semântica e segmentação por instância em imagens macroscópicas placentárias. As redes selecionadas apresentam características complementares: modelos clássicos e robustos, baseados em conexões residuais (**ResNet**), foram comparados a arquiteturas

modernas, focadas em alta eficiência e menor número de parâmetros (**EfficientNet**). Adicionalmente, foi incluída a **YOLOv11**, uma rede moderna de estágio único, para a tarefa complementar de detecção e segmentação por instâncias.

As arquiteturas avaliadas foram:

- a) **YOLO**: Uma arquitetura de estágio único para detecção e segmentação de instâncias, conhecida por sua velocidade e eficácia. Foi incluída para comparação com a capacidade de detecção de objetos (KHANAM; HUSSAIN, 2024; JOCHER; QIU; CHAURASIA, 2024).
- b) **Variante ResNet (ResNet18, ResNet34, ResNet50, ResNet101, ResNet152) com U-Net**: Redes neurais convolucionais profundas que utilizam conexões residuais para mitigar o problema do desaparecimento do gradiente, permitindo o treinamento de modelos com muitas camadas. As diferentes profundidades (18 a 152 camadas) permitem investigar o impacto da complexidade do modelo (ESTEVA *et al.*, 2021).
- c) **U-Net com variantes EfficientNet (EfficientNet-B0, B1, B2, B3, B4, B5, B6, B7)**: Modelos que utilizam um método de escalonamento composto para otimizar a profundidade, largura e resolução da rede de forma equilibrada, alcançando alta precisão com menor número de parâmetros. A gama completa de B0 a B7 foi avaliada para identificar o ponto ideal de desempenho e eficiência (ESTEVA *et al.*, 2021; YU, 2021).

### 4.2.3 Configurações e Estratégias de Treinamento

Os principais hiperparâmetros e estratégias de treinamento empregados para os modelos YOLO, ResNet e EfficientNet são descritos a seguir, e estão resumidos na Tabela 3.

- **Estratégia de Treinamento**: Foi adotada a técnica de Transfer Learning combinada com redimensionamento progressivo (progressive resizing) em três estágios para os modelos ResNet e EfficientNet (U-Net). Esta abordagem permite à rede aprender características gerais em resoluções menores antes de refinar detalhes e texturas na resolução total. O modelo YOLO utilizou um Treinamento Contínuo.
- **Otimizador e Política de LR**: O otimizador AdamW foi empregado em todas as arquiteturas. Nos modelos ResNet e EfficientNet, esta otimização foi acoplada à política de Superconvergência 1Cycle (SMITH, 2017) para maximizar a velocidade de convergência e a capacidade de generalização.
- **Tamanhos de Imagem**: As resoluções de entrada foram ajustadas para cada modelo:  $1600 \times 1200$  (ResNet - máxima),  $512 \times 512$  (EfficientNet - máxima) e  $640 \times 640$  (YOLO).

Tabela 3 – Configuração de treinamento para cada arquitetura avaliada.

| CNN          | ARQUITETURA   MODELO PRÉ-TREINADO | BACKBONE        | ESTRATÉGIA DE TREINO  | TAMANHO IMAGEM ORIGINAL | BATCH SIZE | ÉPOCAS              | OTIMIZADOR | Nº DE IMAGENS (TRAIN   VAL   TEST) |
|--------------|-----------------------------------|-----------------|---|-------------------------|------------|---------------------|------------|------------------------------------|
| YOLO         | YOLOV11                           | YOLO11n-seg     | Continuous Training   | 640 x 640               | 64         | 200                 | AdamW      | 318   33   17                      |
| RESNET       | U-NET                             | ResNet18        | One Cycle Policy + LR Finder, 3-stage Transfer learning + Fine-tuning | 1600 x 1200             | 8   4   2  | 50 + 50 Fine-tuning | AdamW      | 375   47   47                      |
|              |                                   | ResNet34        |   |                         |            |                     |            |                                    |
|              |                                   | ResNet50        |   |                         |            |                     |            |                                    |
|              |                                   | ResNet101       |   |                         |            |                     |            |                                    |
|              |                                   | ResNet152       |   |                         |            |                     |            |                                    |
| EFFICIENTNET | U-NET                             | EfficientNet-B0 | One Cycle Policy, 3-stage Multiresolution + Fine-tuning               | 512 x 512               | 16         | 50 + 50 Fine-tuning | AdamW      | 377   45   46                      |
|              |                                   | EfficientNet-B1 |   |                         |            |                     |            |                                    |
|              |                                   | EfficientNet-B2 |   |                         |            |                     |            |                                    |
|              |                                   | EfficientNet-B3 |   |                         |            |                     |            |                                    |
|              |                                   | EfficientNet-B4 |   |                         |            |                     |            |                                    |
|              |                                   | EfficientNet-B5 |   |                         |            |                     |            |                                    |
|              |                                   | EfficientNet-B6 |   |                         |            |                     |            |                                    |
|              |                                   | EfficientNet-B7 |   |                         |            |                     |            |                                    |

Fonte: Elaborado pela autora (2025).

### 4.3 MÉTRICAS DE AVALIAÇÃO DE DESEMPENHO

A performance dos modelos foi avaliada com base em métricas amplamente reconhecidas para tarefas de segmentação, visando capturar tanto a qualidade geral quanto o comportamento por classe. Para os modelos de segmentação semântica (U-Net com diferentes backbones), adotaram-se:

- **Dice Coefficient (F1-Score):** quantifica a similaridade entre as máscaras preditas e as reais, sendo especialmente sensível a classes minoritárias (ISENSEE *et al.*, 2021).
- **IoU:** também conhecido como Jaccard Index, avalia a proporção de interseção sobre a união das regiões segmentadas (COSTANZO *et al.*, 2018).
- **Foreground Accuracy:** mede a acurácia excluindo a classe background, destacando o desempenho nas estruturas morfológicas relevantes (BARKER *et al.*, 2010).

Para o modelo de segmentação por instância (YOLOv11), foram utilizadas:

- **AP50:** média da precisão considerando limiar de IoU  $\geq 0.5$ .
- **AP50–95:** média da precisão para múltiplos limiares (0.5 a 0.95, com passo 0.05), oferecendo avaliação mais rigorosa (PADILLA; NETTO; SILVA, E. A. B. da, 2020).

Complementarmente, aplicou-se:

- **Curvas de perda (train/val loss)** ao longo das épocas, úteis para monitorar aprendizado, estabilidade e overfitting.

- **Matrizes de confusão multiclasse**, analisando confusões frequentes entre classes visualmente similares.
- **Visualizações qualitativas**, com comparações entre máscaras reais e preditas, para análise subjetiva da performance.

A combinação de métricas objetivas e análises visuais proporcionou uma avaliação abrangente do desempenho dos modelos, permitindo comparação criteriosa entre diferentes arquiteturas.

## 5 RESULTADOS

### 5.1 DESEMPENHO GLOBAL

Esta seção apresenta os resultados experimentais obtidos nas tarefas de segmentação semântica e segmentação por instância das estruturas morfológicas da placenta. As análises quantitativas são complementadas por visualizações qualitativas, matrizes de confusão e gráficos de treinamento, permitindo uma avaliação abrangente do comportamento dos modelos em diferentes aspectos do processo de aprendizado.

#### 5.1.1 Métricas de Segmentação Semântica

As redes baseadas em U-Net seguiram uma estratégia de treinamento em três estágios com resoluções crescentes ( $\frac{1}{4}$ ,  $\frac{1}{2}$  e total), o que permitiu melhor refinamento espacial das máscaras preditas. A Tabela 4 resume os principais resultados obtidos nas métricas Dice, IoU, acurácia e tempo de execução (total) para cada arquitetura avaliada.

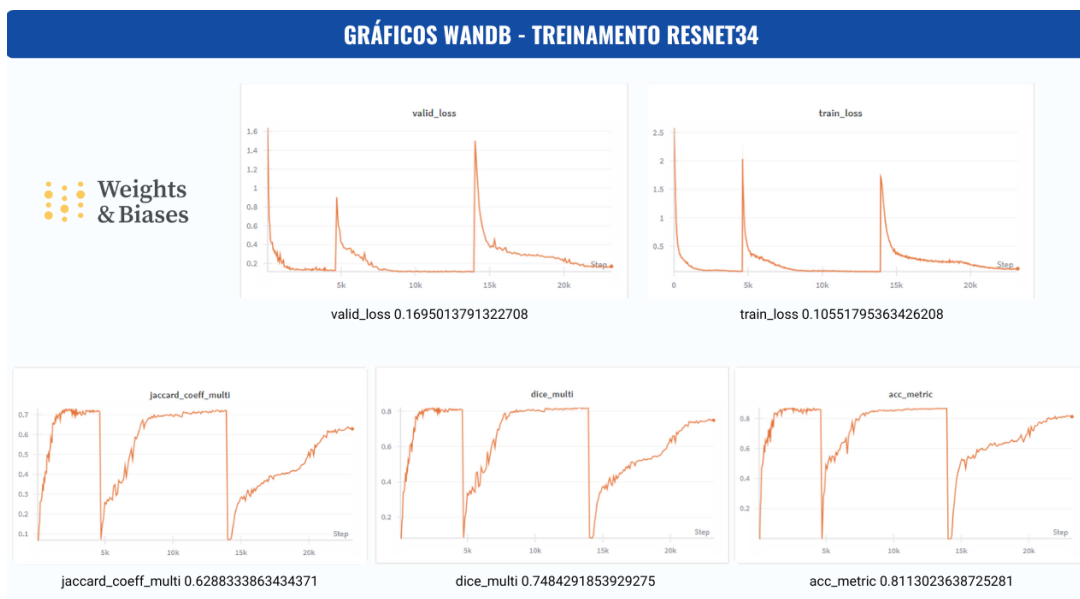
Tabela 4 – Principais métricas de desempenho das arquiteturas U-Net (ResNet e EfficientNet).

| Arquitetura            | Dice          | IoU           | Acurácia      | Tempo de Execução (Total) |
|------------------------|---------------|---------------|---------------|---------------------------|
| <b>ResNet18</b>        | <b>75,54%</b> | <b>63,88%</b> | <b>81,68%</b> | <b>2h 14m 52s</b>         |
| <b>ResNet34</b>        | <b>80,71%</b> | <b>70,01%</b> | <b>85,55%</b> | <b>4h 30m 8s</b>          |
| ResNet50               | 56,05%        | 44,26%        | 68,30%        | 4h 50m 35s                |
| ResNet101              | 65,27%        | 55,05%        | 76,18%        | 6h 50m 45s                |
| ResNet152              | 63,86%        | 52,60%        | 73,52%        | 4h 57m 16s                |
| EfficientNet-B0        | 74,92%        | 68,46%        | 90,10%        | 43m 2s                    |
| EfficientNet-B1        | 74,79%        | 68,35%        | 90,42%        | 43m 41s                   |
| EfficientNet-B2        | 75,48%        | 69,42%        | 91,21%        | 44m 27s                   |
| EfficientNet-B3        | 75,45%        | 69,44%        | 91,20%        | 46m 25s                   |
| <b>EfficientNet-B4</b> | <b>75,69%</b> | <b>69,71%</b> | <b>91,45%</b> | <b>49m 6s</b>             |
| EfficientNet-B5        | 74,24%        | 67,53%        | 89,82%        | 49m 35s                   |
| EfficientNet-B6        | 75,21%        | 68,85%        | 90,46%        | 51m 16s                   |
| <b>EfficientNet-B7</b> | <b>75,78%</b> | <b>69,78%</b> | <b>91,33%</b> | <b>43m 2s</b>             |

Fonte: Elaborado pela autora (2025).

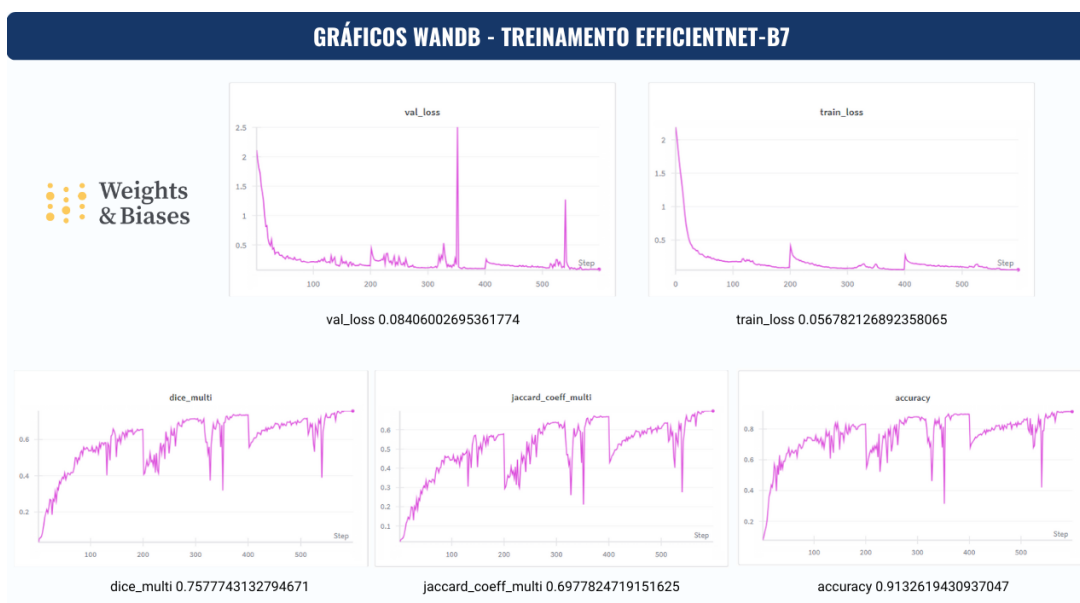
As Figuras 20 e 21 apresentam a evolução temporal da função de perda e das métricas de desempenho para as arquiteturas ResNet34 e EfficientNet-B7. A análise desses gráficos evidencia como cada modelo progride ao longo das épocas, destacando o comportamento da convergência e a correspondência entre os resultados obtidos em treinamento e validação.

Figura 20 – Evolução das métricas de treinamento e validação da arquitetura U-Net ResNet34.



Fonte: Elaborado pela autora (2025).

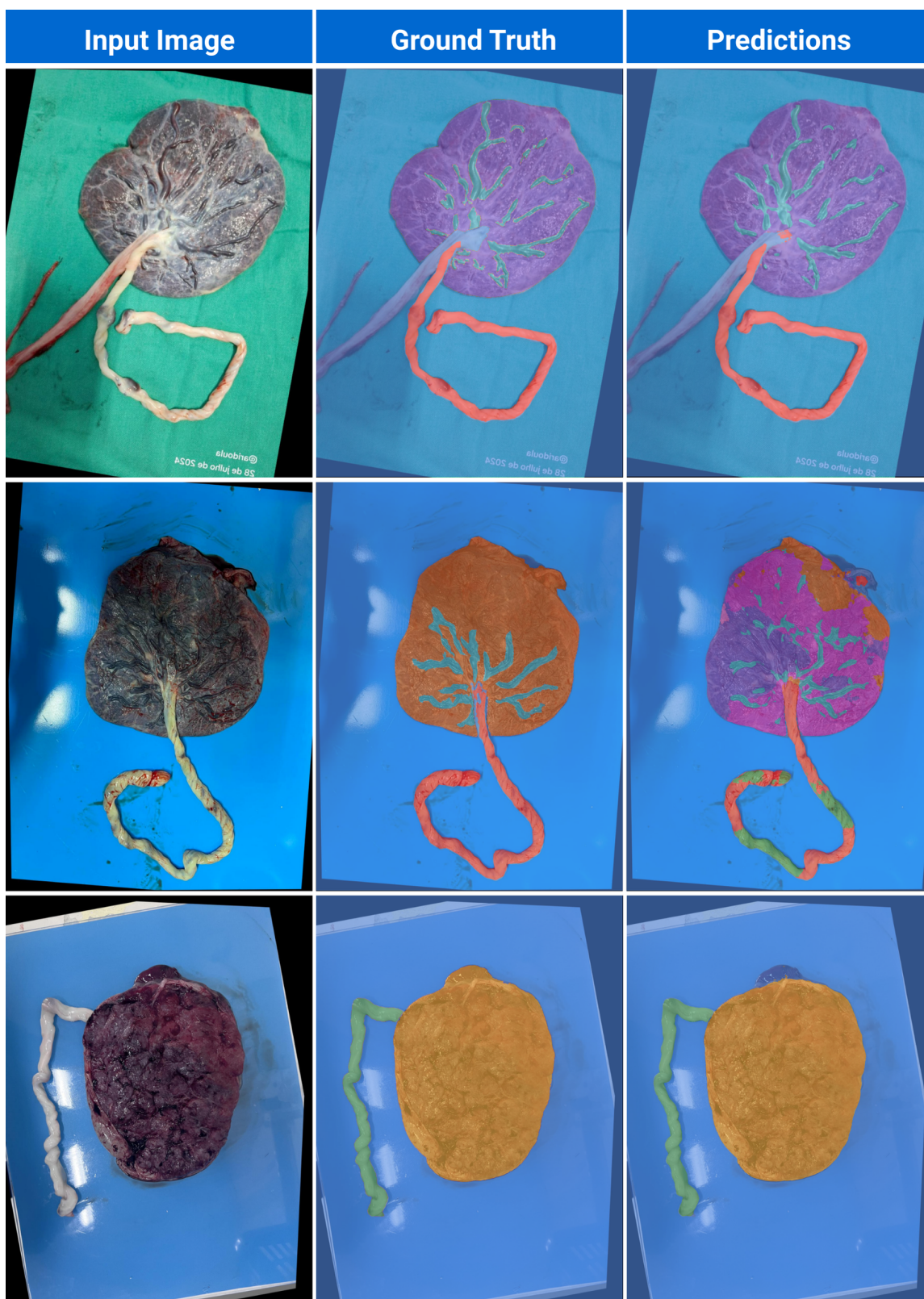
Figura 21 – Evolução das métricas de treinamento e validação da arquitetura U-Net EfficientNet-B7.



Fonte: Elaborado pela autora (2025).

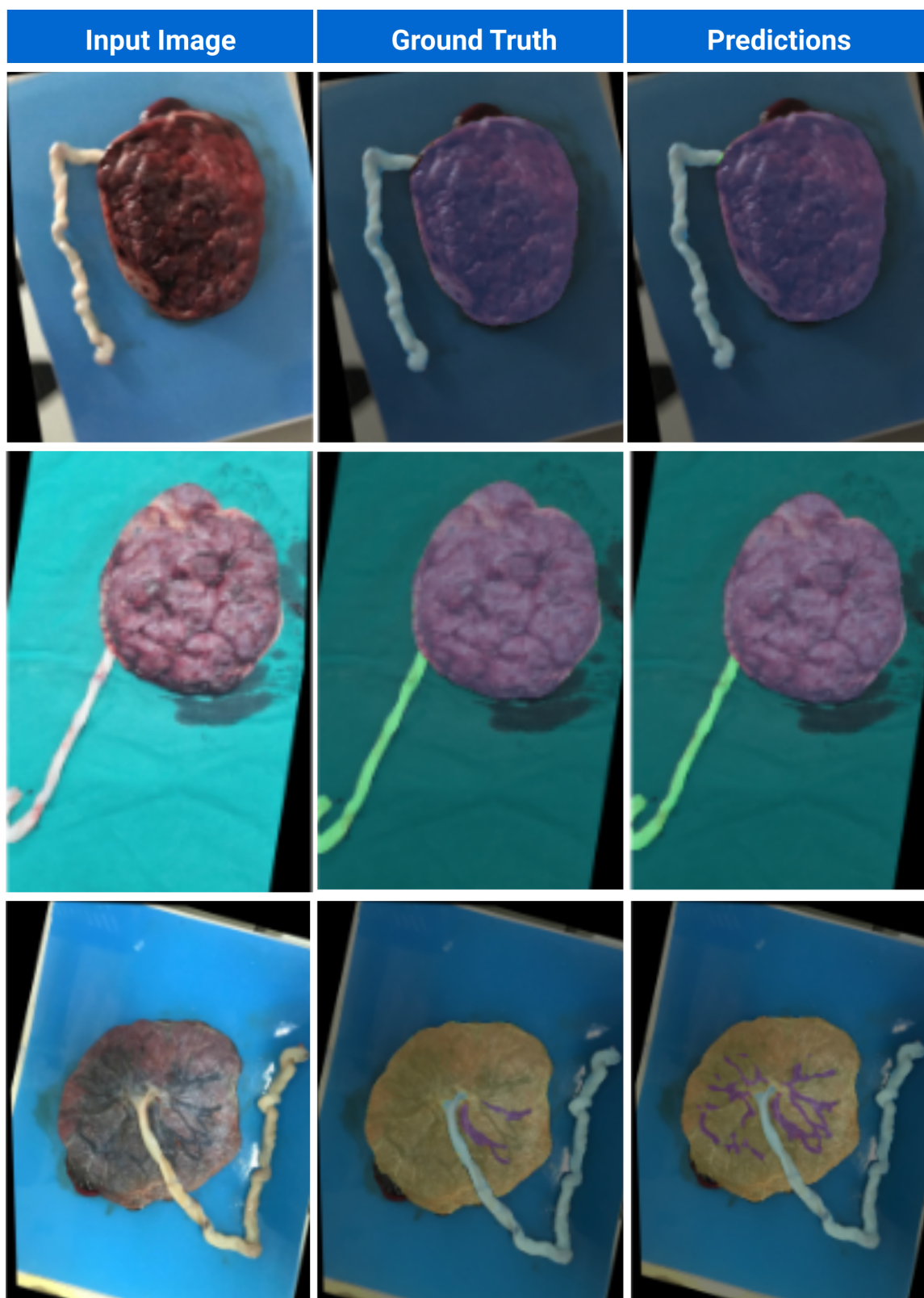
Para ilustrar visualmente o desempenho das arquiteturas, foram selecionados exemplos representativos comparando a imagem de entrada, a anotação manual (ground truth) e a predição gerada pelos modelos. As Figuras 22 e 23 evidenciam a capacidade dos modelos em identificar e segmentar regiões morfológicas distintas da placenta, destacando diferenças visuais no detalhamento das máscaras produzidas.

Figura 22 – Exemplos comparativos entre imagem de entrada, verdade fundamental e predição obtida pela ResNet34.



Fonte: Elaborado pela autora (2025).

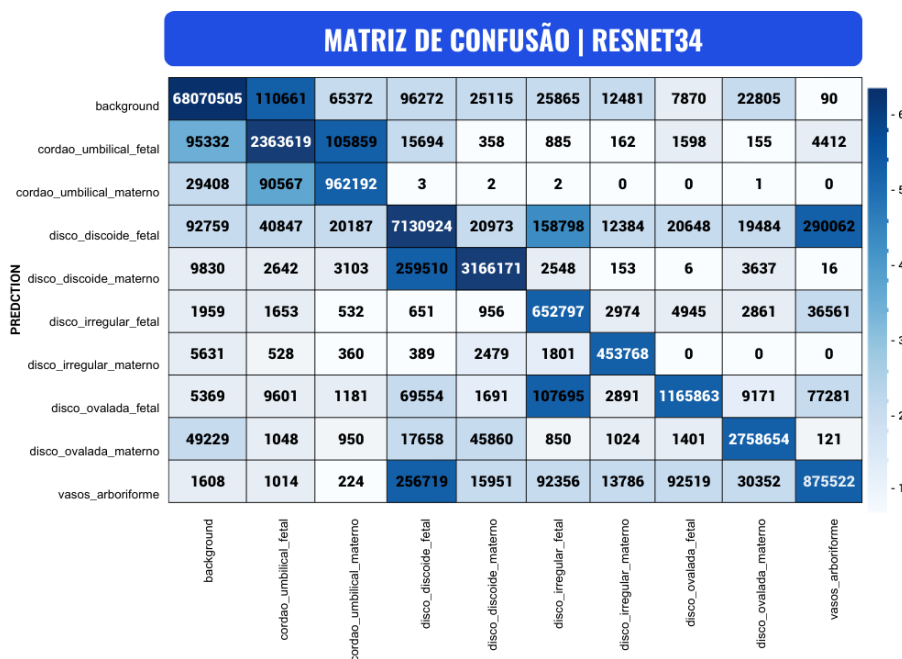
Figura 23 – Exemplos comparativos entre imagem de entrada, verdade fundamental e predição obtida pela EfficientNet-B7.



Fonte: Elaborado pela autora (2025).

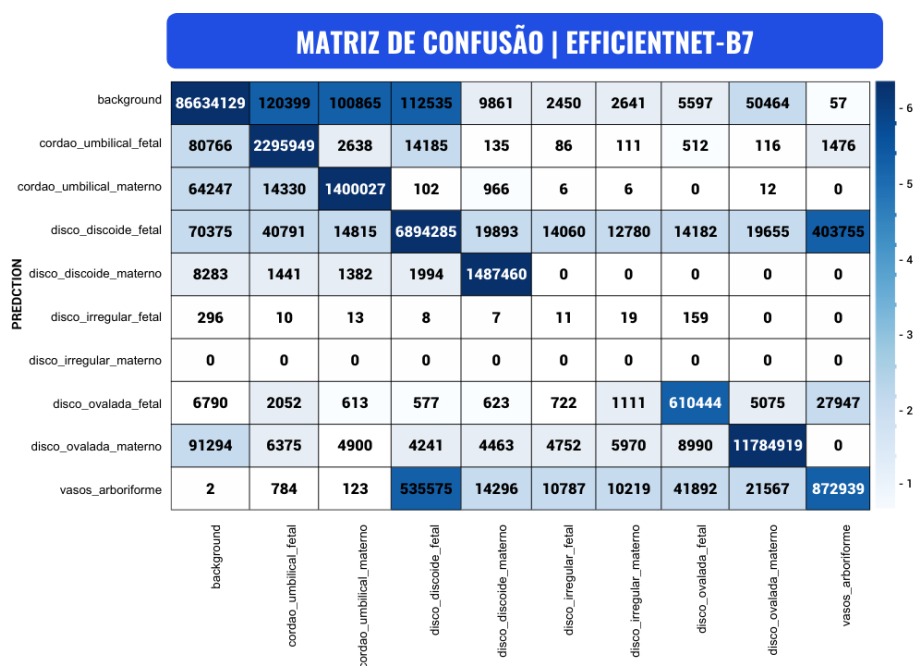
As matrizes de confusão (Figura 24 e 25) geradas para as arquiteturas U-Net permitem observar a distribuição dos acertos e confusões entre as classes segmentadas.

Figura 24 – Matriz de confusão gerada para ResNet34.



Fonte: Elaborado pela autora (2025).

Figura 25 – Matriz de confusão gerada para EfficientNet-B7.












Fonte: Elaborado pela autora (2025).

Para complementar a análise, foram selecionados exemplos de casos nos quais

a segmentação apresentou maior dificuldade, evidenciando limitações específicas observadas durante os experimentos com a ResNet34.

Figura 26 – Exemplos de casos mais desafiadores para a ResNet34.

| Piores Casos  |   |   |
|---|---|---|
| Input Image   | Target  | Predictions   |
|    |    | Loss = 0.1921<br>   |
|   |   | Loss = 0.1572<br>  |
|  |  | Loss = 0.1543<br> |

Fonte: Elaborado pela autora (2025).

### 5.1.2 Métricas de Segmentação por Instância

A arquitetura YOLO foi empregada inicialmente na tarefa de detecção (caixas delimitadoras) e posteriormente na tarefa de segmentação por instância (masks). A

Tabela 5 apresenta os valores de Precision, Recall, mAP@50 e mAP@50–95 para ambas as modalidades.

Tabela 5 – Métricas de detecção (box) e segmentação por instância (mask) obtidas com o modelo YOLO.

| YOLO11 | Precision | Recall | mAP50 | mAP50-95 | Tempo de Execução (Total) |
|--------|-----------|--------|-------|----------|---------------------------|
| Box    | 52,6%     | 72,6%  | 70,3% | 61,5%    | 40min 16s                 |
| Mask   | 50,7%     | 71,4%  | 69,5% | 52,8%    |                           |

Fonte: Elaborado pela autora (2025).

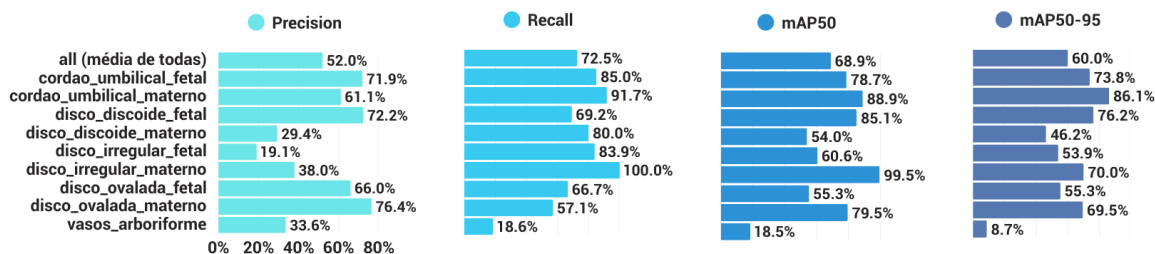
Os resultados demonstram que o modelo obteve desempenho consistente, especialmente nas classes com maior distinção visual, como cordão umbilical e discos placentários, que apresentaram os maiores valores de precisão média. Por outro lado, a classe vasos arboriformes apresentou desempenho inferior, refletindo a alta complexidade morfológica e o desbalanceamento presente no conjunto de dados.

O gráfico das métricas por classe (Figura 27) evidencia com clareza essas diferenças de desempenho entre as estruturas. Observa-se que classes com limites bem definidos apresentam maior precisão, enquanto estruturas ramificadas ou de textura complexa exibem maior dispersão nas medidas de recall e precisão.

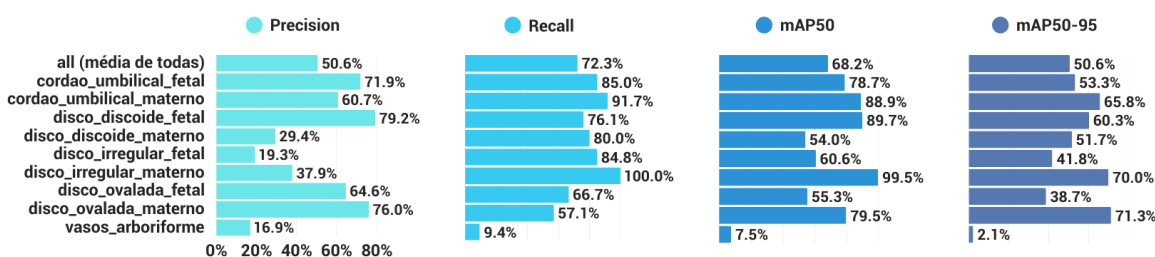
A visualização qualitativa dos resultados reforça esses padrões: em imagens com elevado contraste e boa iluminação, a YOLOv11 identifica e segmenta adequadamente as regiões de interesse; porém, em situações de presença excessiva de fluidos biológicos ou oclusões, verifica-se maior ocorrência de falsos negativos. A Figura 29 contendo exemplos de anotações e predições ilustra essa variação de desempenho.

Figura 27 – Métricas por classe (detecção e segmentação por instância) obtidas com a YOLO.

**Métricas Detecção (Box)**

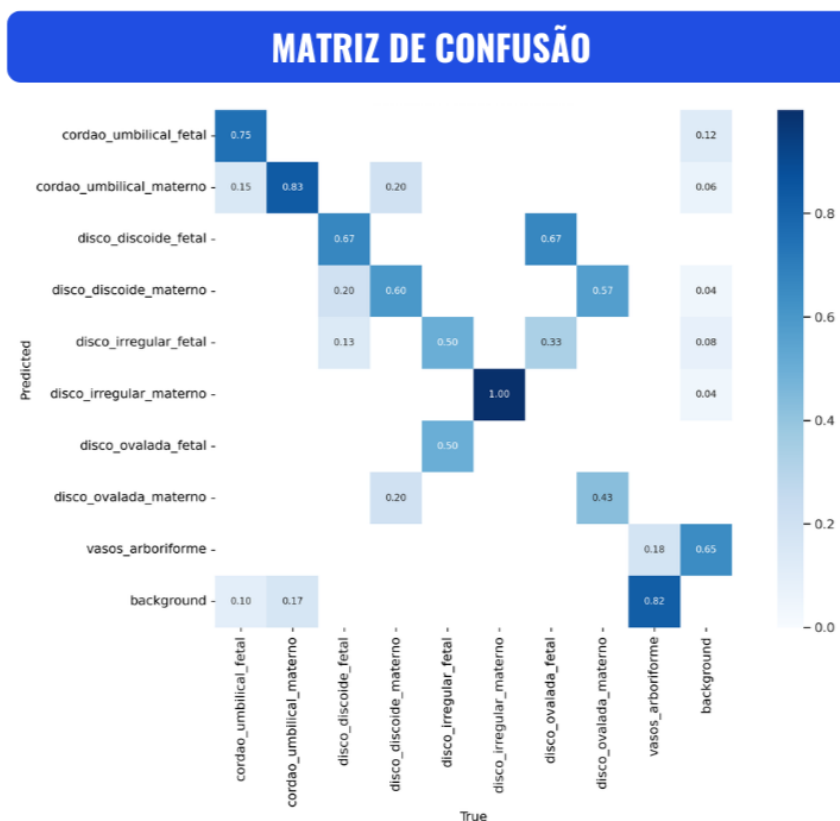


**Métricas Segmentação (Mask)**



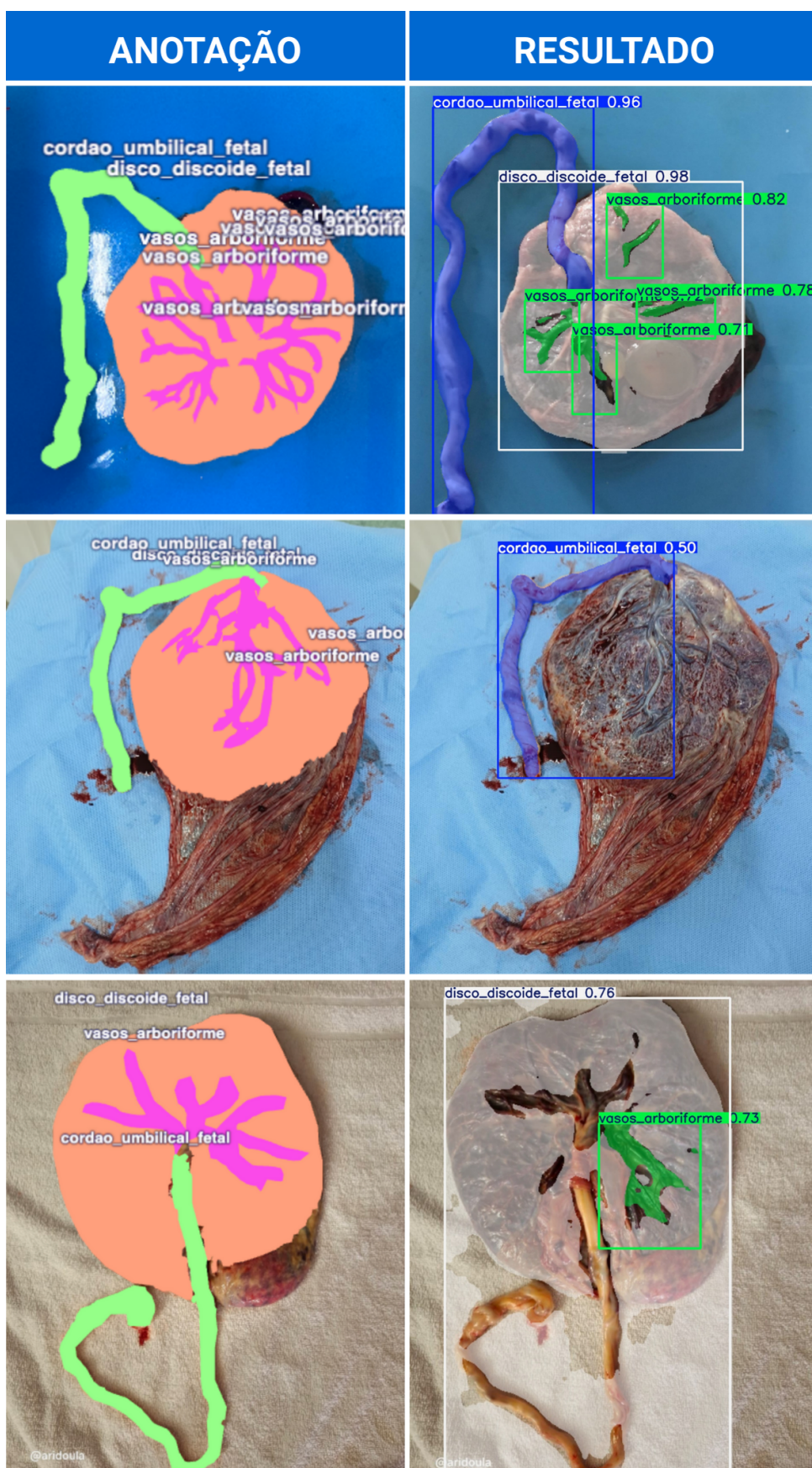
Fonte: Elaborado pela autora (2025).

Figura 28 – Matriz de confusão normalizada do modelo YOLOv11 nas tarefas de detecção e segmentação.



Fonte: Elaborado pela autora (2025).

Figura 29 – Exemplos de anotações manuais e predições geradas pela YOLO.



Fonte: Elaborado pela autora (2025).

## 6 CONSIDERAÇÕES FINAIS

### 6.1 DISCUSSÃO

O presente estudo investigou a aplicação de técnicas de DL para segmentação semântica e segmentação por instância de estruturas morfológicas da placenta em imagens macroscópicas capturadas diretamente na sala de parto. Em contraste com a maior parte da literatura, que utiliza imagens laboratoriais, ultrassonográficas ou obtidas em ambientes controlados, esta pesquisa empregou fotografias reais de pós-parto, buscando aproximar o modelo computacional da complexidade do cenário clínico. Essa escolha amplia a aplicabilidade prática da solução, mas também introduz desafios substanciais, relacionados principalmente à variabilidade de iluminação, presença de sangue, contraste irregular, textura heterogênea e diferenças anatômicas expressivas entre as classes.

O protocolo fotográfico padronizado assumiu papel central na mitigação dessas dificuldades. As imagens obtidas seguindo esse protocolo exibiram maior uniformidade de cor e contraste, favorecendo tanto o processo de anotação quanto o aprendizado automático. A Figura 29 (não exibida aqui, mas mencionada na dissertação) ilustra de modo claro essa diferença: enquanto na imagem padronizada o modelo consegue reconhecer múltiplas estruturas, nas fotografias sem padronização a identificação torna-se substancialmente mais difícil, resultando em falhas de detecção e segmentação. Este efeito reforça a importância de condições controladas de aquisição, mesmo em cenários clínicos não laboratoriais.

No que se refere ao desempenho dos modelos, as arquiteturas U-Net com encoders ResNet34 e EfficientNet-B7 apresentaram os melhores resultados entre as 14 variantes avaliadas. A ResNet34 destacou-se com os valores mais elevados de Dice e IoU, métricas diretamente relacionadas à qualidade da sobreposição entre predição e anotação manual. Seu equilíbrio entre profundidade, estabilidade de treinamento e capacidade discriminativa corrobora achados presentes na literatura de segmentação biomédica, que destacam a adequação de encoders intermediários para bases de dados com variabilidade moderada. Por outro lado, a EfficientNet-B7 obteve a melhor acurácia global entre as arquiteturas avaliadas, além do menor tempo de execução entre as redes U-Net utilizadas, sugerindo maior eficiência computacional quando devidamente ajustada.

A comparação entre ResNet34 e EfficientNet-B7 evidencia que o desempenho não depende exclusivamente da profundidade da arquitetura, mas do alinhamento entre sua capacidade de extração de características e a variabilidade do conjunto de dados. As curvas de treinamento e validação apresentadas ao longo da dissertação confirmam que ambas as arquiteturas alcançaram convergência estável, ainda que a EfficientNet-B7 tenha demonstrado maior sensibilidade aos hiperparâmetros, deman-

dando ajustes mais refinados para atingir desempenho ótimo.

Na segmentação por instância, a YOLOv11 apresentou resultados consistentes tanto para detecção quanto para a geração de máscaras, com métricas superiores nas classes de anatomia mais regular, como cordão umbilical e discos placentários. O desempenho reduzido observado para vasos arboriformes indica a dificuldade intrínseca associada à segmentação de estruturas ramificadas, finas e de baixo contraste. A matriz de confusão e as métricas por classe reforçam essa característica, revelando maior taxa de erros em estruturas visualmente semelhantes ou parcialmente ocluídas. Esse padrão é coerente com estudos prévios na literatura, que apontam a necessidade de abordagens especializadas ou modelos híbridos para lidar com estruturas vasculares complexas.

As análises qualitativas complementaram as evidências quantitativas, permitindo observar, por meio das predições exemplificadas, tanto casos de alto desempenho quanto situações de falha significativa. O conjunto de “piores casos” da ResNet34 permitiu identificar padrões de erro recorrentes, frequentemente associados à ausência de padronização na captura das imagens. Tais achados reforçam a importância de diretrizes claras de aquisição e sugerem que a adoção de protocolos simples pode melhorar substancialmente a qualidade das segmentações produzidas.

A despeito dos bons resultados alcançados, o estudo também evidenciou limitações oriundas da dimensão reduzida do dataset e do desbalanceamento entre classes. Estruturas menos representadas apresentaram menor desempenho, mesmo após ajustes de treinamento, o que é esperado em bases heterogêneas. Esses aspectos serão aprofundados na subseção seguinte sobre ameaças à validade.

Por fim, cabe destacar que este trabalho avança na literatura ao demonstrar a viabilidade de segmentação multiclasse da placenta em imagens reais de sala de parto — um domínio ainda escasso em estudos científicos. A integração entre estratégia de multirresolução, protocolo de aquisição padronizado e avaliação multicritério contribui para consolidar a base metodológica necessária ao desenvolvimento de ferramentas automatizadas aplicáveis ao contexto clínico. Embora não tenha sido realizada validação em campo, os resultados obtidos indicam que soluções assistivas baseadas em inteligência artificial, como as idealizadas nos projetos AI-plax e PlacentaNet, possuem potencial para auxiliar a triagem pós-parto, otimizando a interpretação inicial da placenta por profissionais da saúde.

### 6.1.1 Ameaças à Validade

A principal ameaça à validade deste estudo refere-se ao tamanho reduzido e à limitação clínica do conjunto de dados utilizado. Embora as imagens tenham sido obtidas com protocolo padronizado e representem condições reais de sala de parto, o volume ainda é pequeno para garantir a generalização dos modelos, especialmente

em cenários clínicos mais variados. Além disso, o desbalanceamento entre classes anatômicas pode ter impactado o aprendizado de certas estruturas. Assim, reforça-se a necessidade de expansão da base de dados com maior diversidade de casos, ampliando a robustez dos modelos e sua aplicabilidade clínica.

## 6.2 CONCLUSÃO

Este trabalho investigou, de forma abrangente, a aplicação de arquiteturas de deep learning para segmentação semântica e segmentação por instância de estruturas morfológicas da placenta em imagens macroscópicas capturadas em ambiente clínico real. Ao longo do estudo, foram analisadas quatorze arquiteturas distintas, incluindo variantes da U-Net com diferentes encoders e o modelo YOLOv11, com o objetivo de estabelecer um benchmark inicial de desempenho e avaliar o potencial dessas abordagens como ferramenta de apoio à triagem automatizada.

Os resultados demonstraram que a U-Net ResNet34 apresentou o melhor desempenho global, alcançando os maiores valores de Dice e IoU, indicando elevada capacidade de representar com precisão a geometria das estruturas segmentadas. Já a U-Net EfficientNet-B7 destacou-se pela maior acurácia global e pelo menor tempo de execução, evidenciando grande eficiência computacional. No contexto da segmentação por instância, o modelo YOLOv11 obteve métricas consistentes nas classes com morfologia mais regular, reforçando sua adequação para tarefas de detecção e segmentação rápida.

Com base nesses achados, é possível responder à hipótese formulada nesta dissertação: Sim, é possível desenvolver e validar um modelo computacional de deep learning, baseado em técnicas de segmentação, capaz de oferecer um benchmark de precisão e avançar em direção à robustez necessária para apoiar a triagem automatizada de alterações macroscópicas placentárias em ambientes clínicos reais.

Todavia, convém destacar que tal confirmação é parcial. Embora os modelos tenham demonstrado viabilidade técnica, bom desempenho quantitativo e capacidade de generalização dentro do escopo do dataset disponível, a robustez clínica plena ainda não pode ser assegurada. Essa limitação decorre principalmente do tamanho reduzido e da baixa variabilidade do conjunto de dados, da sensibilidade à padronização da imagem, e da ausência de validação em campo com profissionais da saúde. Assim, a hipótese se confirma no plano técnico-experimental, mas sua consolidação em ambiente clínico demanda avanços adicionais.

Em síntese, esta pesquisa contribui significativamente para o avanço da literatura sobre segmentação placentária, introduzindo um pipeline metodologicamente sólido, demonstrando a aplicabilidade de modelos modernos em imagens reais e estabelecendo bases concretas para investigações futuras. Os resultados apresentados reforçam o potencial da inteligência artificial como ferramenta de apoio à triagem

pós-parto e oferecem diretrizes importantes para o desenvolvimento de soluções mais robustas, escaláveis e integradas ao fluxo de trabalho clínico.

### 6.2.1 Trabalhos Futuros

A partir dos resultados obtidos e das limitações identificadas ao longo deste estudo, sugerem-se as seguintes direções para pesquisas futuras:

#### 1. Expansão e diversificação do dataset

- Aumentar a quantidade de imagens, incluindo dados provenientes de diferentes maternidades, equipamentos e populações.
- Ampliar a variabilidade clínica e de condições de captura, de modo a fortalecer a capacidade de generalização dos modelos.

#### 2. Refinamento das anotações

- Reanotar estruturas complexas, como vasos arboriformes, com maior nível de detalhe e consistência.
- Implementar revisão cruzada entre especialistas para reduzir discrepâncias e assegurar qualidade nas máscaras manuais.

#### 3. Avaliação de modelos baseados em atenção

- Investigar arquiteturas como Swin-UNet, SegFormer e TransUNet, que apresentam desempenho superior em estruturas irregulares e de baixa definição.

#### 4. Técnicas avançadas de aumento e normalização de dados

- Desenvolver estratégias de *data augmentation* específicas para simular variabilidade luminosa, presença de fluidos e artefatos comuns em sala de parto.
- Explorar métodos de *domain adaptation* e normalização de cor para reduzir a dependência da padronização na captura das imagens.

#### 5. Validação clínica e análise de usabilidade

- Testar os modelos em fluxos reais de atendimento, avaliando desempenho em ambiente clínico.
- Analisar interpretabilidade, confiabilidade e aceitação por parte de profissionais da saúde.

#### 6. Desenvolvimento de um sistema assistivo

- Criar protótipos de aplicações móveis ou web que permitam inferência em tempo real no ponto de cuidado.

- Investigar técnicas de compressão de modelos, como poda, quantização e distilação, para viabilizar execução eficiente em hardware de baixo custo.

## REFERÊNCIAS

- BARKER, David J. P.; OSMOND, Clive; THORNBURG, Kent L.; KAJANTIE, Eero; ERIKSSON, Johan G. The prenatal origins of lung cancer. II. The placenta. **Placenta**, Elsevier, v. 31, n. 7, p. 568–572, jul. 2010.
- BIEWALD, L. Experiment Tracking with Weights and Biases. **M.S. Thesis, Stanford University**, 2020.
- BUSLAEV, A.; PARINOV, A.; KHVEDCHENYA, E.; IGLOVIKOV, V. I.; KALININ, A. A. Alumentations: Fast and flexible image augmentations. **Information**, v. 11, n. 2, p. 125, 2020.
- CHEN, Y.; WU, C.; ZHANG, Z.; GOLDSTEIN, J. A.; GERNAND, A. D.; WANG, J. Z. Automatic morphological characterization of placenta photos with deep learning. *In*: PROC. Med. Image Comput. Comput. Assist. Interv. (MICCAI). Cham, Switzerland: Springer, 2019. v. 11764, p. 487–495.
- CHEN, Y.; ZHANG, Z.; WU, C.; DAVAASUREN, D.; GOLDSTEIN, J. A.; GERNAND, A. D.; WANG, J. Z. Ai-plax: Ai-based placental assessment and examination using photos. **Computerized Medical Imaging and Graphics**, v. 84, p. 101744, 2020.
- COSTANZO, Vincenzo; BARDELLI, Alberto; SIENA, Salvatore; ABRIGNANI, Sergio. Exploring the links between cancer and placenta development. **Open Biology**, The Royal Society, v. 8, n. 6, p. 180081, jun. 2018.
- DWYER, B. **Understanding Auto Orientation and EXIF Data**. [S.l.: s.n.], 2020. Roboflow Blog. [Online; Acesso em <2025>]. Disponível em: <https://blog.roboflow.com/exif-auto-orientation/>.
- ESTEVA, Andre *et al.* Deep learning-enabled medical computer vision. **npj Digital Medicine**, Nature Publishing Group, v. 4, n. 1, p. 5, jan. 2021.
- GIL, A. C. **Como Elaborar Projetos de Pesquisa**. 4. ed. São Paulo: Editora Atlas S.A., 2002.
- HOWARD, J.; GUGGER, S. Fastai: A Layered API for Deep Learning. **MDPI Information**, v. 11, n. 2, p. 108, 2020.
- HUNTER, J. D. Matplotlib: A 2D Graphics Environment. **Computing In Science & Engineering**, v. 9, n. 3, p. 90–95, 2007.
- ISENSEE, Fabian; JAEGER, Paul F.; KOHL, Simon A. A.; PETERSEN, Jens; MAIER-HEIN, Klaus H. nnU-Net: a self-configuring method for deep learning-based

biomedical image segmentation. **Nature Methods**, Nature Publishing Group, v. 18, p. 203–211, fev. 2021.

JAKHAR, D.; KAUR, I. Artificial intelligence, machine learning and deep learning: definitions and differences. **Clinical and Experimental Dermatology**, v. 45, n. 1, p. 131–132, 2020.

JOCHER, G.; QIU, J.; CHAURASIA, A. **Ultralytics YOLO 11**. [S.l.: s.n.], 2024. GitHub Repository. Acesso em: [Inserir Data de Acesso]. Disponível em: <https://github.com/ultralytics/ultralytics>.

KHANAM, R.; HUSSAIN, M. Yolov11: An overview of the key architectural enhancements. **arXiv preprint arXiv:2410.17725**, 2024. arXiv: 2410.17725 [cs.CV].

KHONG, T. Y. *et al.* Sampling and definitions of placental lesions: Amsterdam placental workshop group consensus statement. **Archives of Pathology & Laboratory Medicine**, v. 140, p. 698–713, 2016.

KITCHENHAM, B. A. **Procedures for Performing Systematic Reviews**. [S.l.], 2004. Disponível em: <https://www.inf.ufsc.br/~5Ctextasciitilde%7B%7Daldo.vw/kitchenham.pdf>.

LAPIX, INCOD, UFSC. **LAPIXDL: Medical Image Segmentation and Visualization Tools**. [S.l.: s.n.], 2025. Software and Documentation, UFSC.

LE CUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.

LECUN, Yann; BOTTOU, Léon; BENGIO, Yoshua; HAFFNER, Patrick. Gradient-Based Learning Applied to Document Recognition. *In*: 11. PROCEEDINGS of the IEEE. [S.l.: s.n.], 1998. v. 86, p. 2278–2324.

LOPES, B. S.; SILVA, B. Q. T.; WANGENHEIM, A. von. **PLACENTA IMAGE CAPTURE PROTOCOL IN THE DELIVERY ROOM / PROTOCOLO DE CAPTURA DE IMAGENS DE PLACENTAS NA SALA DE PARTO**. Florianópolis, 2025.

LOPES, B. S.; SILVA, B. Q. T.; WANGENHEIM, A. von. **Systematic Literature Review on Computer Vision and Deep Learning for Automated Screening of Macroscopic Placental Alterations**. Florianópolis, 2025.

MCKINNEY, W. Data Structures for Statistical Computing in Python. **Proceedings of the 9th Python in Science Conference**, p. 51–56, 2010.

NOGUEIRA, D. S. **EfficientUNet-PyTorch: UNet with EfficientNet backbone**. [S.l.: s.n.], 2021. GitHub Repository.

- PADILLA, Rafael; NETTO, Sergio L.; SILVA, Eduardo A. B. da. A Survey on Performance Metrics for Object-Detection Algorithms. *In: 2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. [S.l.]: IEEE, 2020. p. 237–242.
- PAN, Y.; CAI, T.; MEHTA, M.; GERNAND, A. D.; GOLDSTEIN, J. A.; MITHAL, L. B.; MWINYELLE, D.; GALLAGHER, K.; WANG, J. Z. Enhancing automatic placenta analysis through distributional feature recomposition in vision language contrastive learning. *In: PROC. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*. [S.l.: s.n.], 2023. v. 14393. (Lecture Notes in Computer Science), p. 116–126.
- PAN, Y.; GERNAND, A. D.; GOLDSTEIN, J. A.; MITHAL, L. B.; MWINYELLE, D.; WANG, J. Z. Vision-language contrastive learning approach to robust automatic placenta analysis using photographic images. *In: PROC. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*. [S.l.: s.n.], 2022. v. 13436. (Lecture Notes in Computer Science), p. 707–716.
- PASZKE, A. *et al.* PyTorch: An imperative style, high-performance deep learning library. **Advances in Neural Information Processing Systems**, 2019.
- PTACEK, I.; SEBIRE, N. J.; MAN, J. A.; BROWNBILL, P.; HEAZELL, A. E. Systematic review of placental pathology reported in association with stillbirth. **Placenta**, v. 35, n. 8, p. 552–562, 2014.
- PURI, Pranav; YIANNIAS, James A.; MANGOLD, Aaron R.; SWANSON, David L.; PITTELKOW, Mark R. The policy dimensions, regulatory landscape, and market characteristics of teledermatology in the United States. **JAAD International**, v. 1, n. 2, p. 202–207, dez. 2020. ISSN 2666-3287.
- ROBERTS, D. J. Placental pathology, a survival guide. **Archives of Pathology & Laboratory Medicine**, v. 132, p. 641–651, 2008.
- SAMUEL, Arthur L. Some studies in machine learning using the game of checkers. **IBM Journal of Research and Development**, v. 44, n. 1.2, p. 206–226, 2000. Reimpressão de: IBM J. Res. Dev. 3(3):210–229 (1959).
- SILVA, B. Q. T.; LOPES, B. S.; SANTOS, L. O.; RIBEIRO, R. de P. e S.; WANGENHEIM, A. von. Towards Automated Placental Screening: Instance Segmentation in Clinical Images. *In: PROCEEDINGS of the 38th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*. Madri, Espanha: [s.n.], 2025.
- SILVA, T. *et al.* A importância do exame anatomopatológico da placenta em obstetria e neonatologia. **HU Revista**, Juiz de Fora, v. 42, n. 3, p. 171–175, 2016.

SMITH, Leslie N. Super-Convergence: Very Fast Training of Residual Networks Using Large Learning Rates. **arXiv preprint arXiv:1708.07120**, 2017.

WANGENHEIM, Aldo von; COMUNELLO, Eros; RICHA, Rogério. **Visão Computacional: Seminário Introdução à Visão Computacional**. [S.l.: s.n.], 2013. The Cyclops Project. PPGCC-INE-UFSC. Acessado em 7 out. 2021. Disponível em: <https://www.inf.ufsc.br/~aldo.vw/visao/>.

WASKOM, M. L. Seaborn: Statistical data visualization. **Journal of Open Source Software**, v. 6, n. 60, p. 3021, 2021.

YU, Zhen. Pneumonia Detection with U-EfficientNet. *In*: 2021 IEEE Sixth International Conference on Data Science in Cyberspace (DSC). Shenzhen, China: IEEE, 2021. p. 591–594.

ZHANG, Wenjun J.; YANG, Guosheng; LIN, Yingzi; JI, Chunli; GUPTA, Madan. On Definition of Deep Learning. *In*: 2018 World Automation Congress (WAC). [S.l.: s.n.], 2018. p. 1–5.

ZHANG, Z.; DAVAASUREN, D.; WU, C.; GOLDSTEIN, J. A.; GERNAND, A. D.; WANG, J. Z. Multi-region saliency-aware learning for cross-domain placenta image segmentation. **Pattern Recognit. Lett.**, v. 140, p. 165–171, 2020.

**APÊNDICE A – SYSTEMATIC LITERATURE REVIEW ON COMPUTER VISION  
AND DEEP LEARNING FOR AUTOMATED SCREENING OF MACROSCOPIC  
PLACENTAL ALTERATION**

**Autores:** Beatriz Silva Lopes, Bibiana Quatrin Tiellet da Silva e Aldo von Wangenheim.

**DOI:** 10.13140/RG.2.2.23456.70405

**Report number:** INCoD/LAPIX.03.2025.E



Instituto Nacional para Convergência Digital



UNIVERSIDADE FEDERAL  
DE SANTA CATARINA

ISSN 2236-5281

Technical Report

**INCoD/LAPIX.03.2025.E**

# **Systematic Literature Review on Computer Vision and Deep Learning for Automated Screening of Macroscopic Placental Alterations**

## **Authors:**

Beatriz Silva Lopes

Bibiana Quatrin Tiellet da Silva

Aldo von Wangenheim

**DOI: 10.13140/RG.2.2.23456.70405**

Version 1.0

Status: Final

Distribution: External

**FEBRUARY - 2025**

© 2025 **INCoD** – Brazilian Institute for Digital Convergence

All rights reserved and protected under Brazilian Law No. 9.610 from 19/02/1998. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise.



**Brazilian Institute for Digital Convergence**

Federal University of Santa Catarina - UFSC  
João David Ferreira Lima University Campus - Trindade  
Department of Informatics and Statistics - Room  
320 Florianópolis-SC - CEP 88040-970

Phone/ FAX: +55 48 3721-9516 R.17

[www.incod.ufsc.br](http://www.incod.ufsc.br)

ISSN 2236-5281

Technical Report from the Brazilian institute for Digital Convergence/ Department of Informatics and Statistics, Technological Center, Federal university of Santa Catarina. -- v.1, n.1 (2024).-- Florianópolis: INE, UFSC, 2024-

Summary in English

ISSN 2236-5281

DOI: 10.13140/RG.2.2.23456.70405

1. Convergência digital. 2. Padronização de imagens. 3. Tecnologia em saúde materna. 4. Universidade Federal de Santa Catarina. Departamento de Informática e Estatística.

# CONTENTS

|   |           |
|---|-----------|
| <b>1 INTRODUCTION.....</b>                      | <b>4</b>  |
| <b>2 SYSTEMATIC LITERATURE REVIEW.....</b>      | <b>5</b>  |
| 2.1 STATE OF THE ART.....                       | 5         |
| 2.1.1 Objective of the Review.....              | 6         |
| 2.1.2 Search Strategy.....                      | 7         |
| 2.1.2.1 Search Terms.....                       | 7         |
| 2.1.2.2 Selection Process.....                  | 10        |
| <b>3 DETAILED ANALYSIS OF THE ARTICLES.....</b> | <b>13</b> |
| 3.1 2019.....                                   | 13        |
| 3.2 2020.....                                   | 15        |
| 3.3 2022.....                                   | 19        |
| 3.4 2023.....                                   | 20        |
| <b>4 DISCUSSION.....</b>                        | <b>23</b> |
| <b>5 CONCLUSION.....</b>                        | <b>27</b> |
| <b>REFERENCES.....</b>                          | <b>28</b> |

# 1 INTRODUCTION

The placenta is a transient organ essential for the exchange of nutrients, oxygen, and metabolites between the mother and fetus, in addition to playing a key role in hormone production and the development of the fetal immune barrier (ROBERTS, 2008). Morphological alterations in the placenta can be associated with gestational complications and adverse perinatal outcomes, accounting for up to 65% of fetal death cases (PTACEK, 2014). Thus, placental analysis is a crucial procedure for investigating pathological events that occur during pregnancy (KHONG, 2016).

Macroscopic placental evaluation allows for the identification of structural abnormalities in the placental disc and umbilical cord, as well as anomalies in coloration and the integrity of the amniotic membranes. In many cases, this analysis determines the need for histopathological examination, which is performed by pathologists upon request from the attending obstetrician (SILVA, 2016). However, the systematic request for placental histopathological examination is unfeasible due to its high cost and operational demands. Therefore, the decision to refer a placenta for laboratory analysis is based on clinical criteria and morphological characteristics observed at the time of delivery. Nevertheless, the subjectivity of macroscopic evaluation and the variability in hospital protocols may impact the early detection of placental pathologies (SILVA, 2016).

The advancement of Artificial Intelligence (AI) in the medical field has driven the development of computational tools for medical image analysis, including prenatal imaging exams such as ultrasound and magnetic resonance imaging (MRI), as well as histological tissue evaluation. However, in the postpartum macroscopic analysis context, there is a notable gap in the literature regarding the development of automated methods for identifying morphological patterns in photographic records of placentas.

Given this scenario, this study presents a Systematic Literature Review (SLR) conducted according to the methodological guidelines proposed by Kitchenham (2004), with the objective of investigating the state of the art in computational methods for macroscopic placental analysis based on photographic images. The review aims to map the techniques used, assess their applicability, and identify challenges that must be overcome to facilitate the adoption of these systems in clinical practice.

## 2 SYSTEMATIC LITERATURE REVIEW

### 2.1 STATE OF THE ART

In order to establish the state of the art in macroscopic placental analysis through computer vision approaches, we conducted a systematic literature review following the protocols proposed by Kitchenham (2004). Figure 1 presents a summary of the structure of our review.

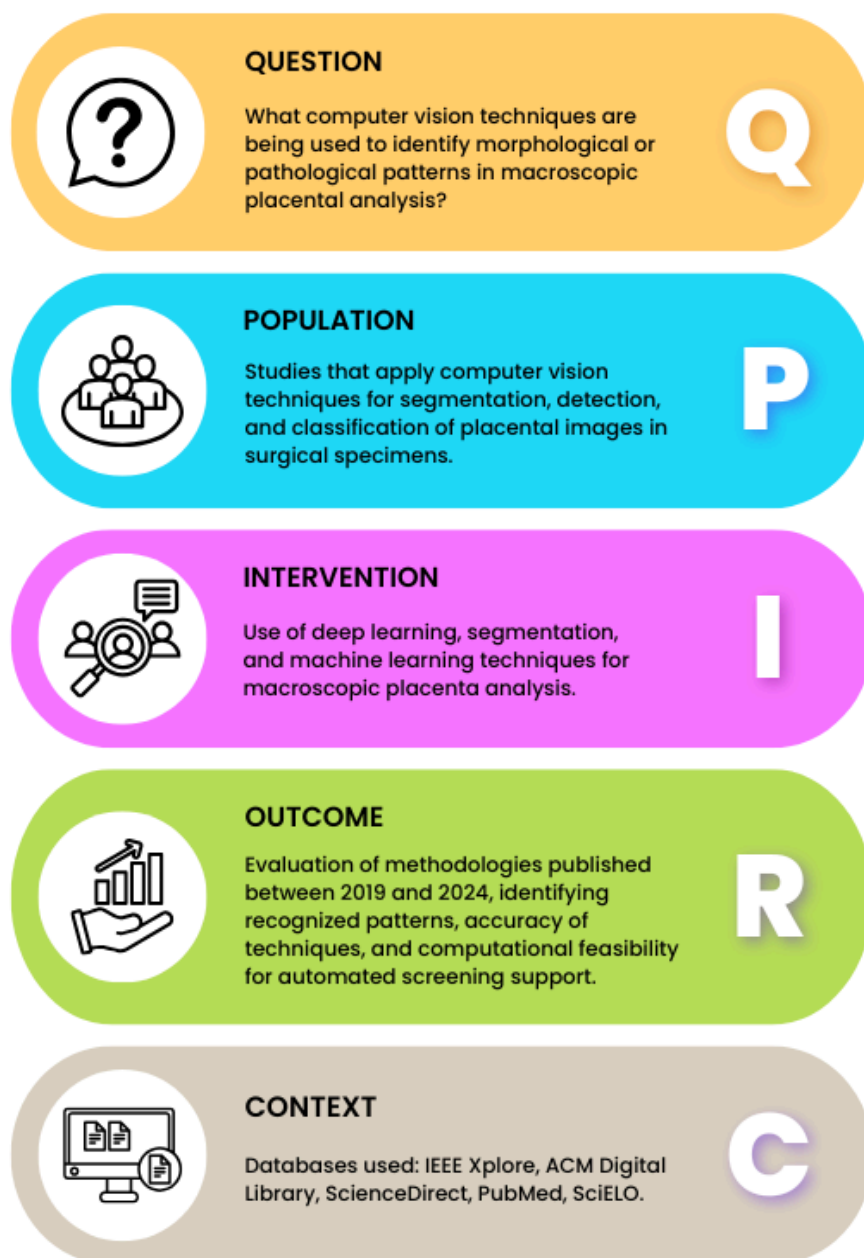


Figure 1. Research Structure. (Source: Authors, 2025).

### 2.1.1 Objective of the Review

The objective of this systematic literature review is to identify, compile, and analyze academic studies related to computer vision methods used in the identification of macroscopic placental alterations in surgical specimens. Based on the research structure presented in Figure 1, we propose the following research questions (Table 2) to be addressed in this systematic review.

|   |
|---|
| <ul style="list-style-type: none"><li>• What techniques are being used for morphological or pathological pattern recognition in macroscopic placental analysis?</li></ul> |
| <ul style="list-style-type: none"><li>• What patterns are identified, and how are they classified?</li></ul>  |
| <ul style="list-style-type: none"><li>• What strategies are used to interpret the shape of the placental disc?</li></ul>  |
| <ul style="list-style-type: none"><li>• What strategies are used to assess the integrity of the placental disc?</li></ul>   |
| <ul style="list-style-type: none"><li>• What strategies are used to analyze the shape and insertion of the umbilical cord?</li></ul>                                      |
| <ul style="list-style-type: none"><li>• What strategies are used to interpret color alterations in the placenta?</li></ul>  |
| <ul style="list-style-type: none"><li>• What contextual variables are incorporated into the model?</li></ul>  |
| <ul style="list-style-type: none"><li>• What are the accuracy, precision, computational cost, and processing mode of these methods?</li></ul>                             |
| <ul style="list-style-type: none"><li>• What is the adaptability and portability of the solution?</li></ul>   |

Table 2. Research Questions.

The results of this review consist of a compilation of methods and techniques used for recognizing placental morphological patterns based on alterations in the organ's shape, integrity, and coloration, with a focus on detecting both normal and pathological morphological patterns. Additionally, the characteristics of the developed applications and their respective contexts are documented. These applications pertain to the use of a model for supporting the automated screening of macroscopic placental alterations and the request for histopathological examination.

## **2.1.2 Search Strategy**

Aiming to identify studies focused on the macroscopic evaluation of placental characteristics in the context of computer vision, and to refine the search terms, an initial informal search was conducted across various platforms, including Google Scholar, IEEE, ACM, and Springer. Based on this preliminary search, the most relevant keywords and their corresponding filters were selected to construct search strings for the databases recommended by Kitchenham, as well as the major medical databases, listed below.

Sources of Information:

- **IEEE Xplore** (<http://ieeexplore.ieee.org>).
- **ScienceDirect** ([www.sciencedirect.com](http://www.sciencedirect.com)).
- **ACM Digital Library** (<http://portal.acm.org>).
- **Pubmed** (<https://pubmed.ncbi.nlm.nih.gov>).
- **Scielo** (<https://www.scielo.br>).

It is worth noting that at this stage, a marked scarcity of studies related to the pathological characterization of macroscopic placental findings was observed, indicating that few studies have been conducted focusing on the evaluation of the organ in the postnatal period. In contrast, there has been a notable increase in studies conducted during the prenatal period, particularly those characterizing morphological findings through radiological imaging exams, such as magnetic resonance imaging (MRI) and ultrasonography.

### **2.1.2.1 Search Terms**

Through a comprehensive review of the studies selected during the informal search, the following keywords, their synonyms, and their respective associations based on proximity relationships were identified:

(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).

Tables 3, 4, 5, 6, and 7 below detail the search strategy applied to each of the databases.

| IEEE Xplore                        |  |                                |   |
|------------------------------------|--|--------------------------------|---|
| <b>Search Terms</b>                | (((("Document Title":"placenta" OR "Abstract":"placenta" OR "Document Title":"pathology" OR "Abstract":"pathology" OR "Document Title":"placental pathology" OR "Abstract":"placental pathology" OR "Document Title":"placental" OR "Abstract":"placental") AND ("Document Title":"deep learning" OR "Abstract":"deep learning" OR "Document Title":"Convolutional neural network" OR "Abstract":"Convolutional neural network" OR "Document Title":"segmentation" OR "Abstract":"segmentation") AND ("Document Title":"medical" OR "Abstract":"medical" OR "Document Title":"health" OR "Abstract":"health" OR "Document Title":"anatomy" OR "Abstract":"anatomy")))) ) |                                |   |
| <b>Refining Search</b>             | <ul style="list-style-type: none"> <li>• 2019 - 2024</li> <li>• Title and Abstract</li> </ul>  |                                |   |
| <b>Quantitative Results</b>        | 287  |                                |   |
| <b>Relevant (Title + Abstract)</b> | 10   | <b>Relevant (Full Article)</b> | 0 |

Table 3. IEEE Xplore Search.

| ACM Digital Library                |  |                                |   |
|------------------------------------|--|--------------------------------|---|
| <b>Search Terms</b>                | [[Title: "placenta"] OR [Title: "placental"] OR [Title: "placental pathology"] OR [Abstract: "placenta"] OR [Abstract: "placental"] OR [Abstract: "placental pathology"]] AND [[Title: "deep learning"] OR [Title: "neural network"] OR [Title: "convolutional neural network"] OR [Title: "cnn"] OR [Title: "segmentation"] OR [Title: "segmentation"] OR [Abstract: "deep learning"] OR [Abstract: "neural network"] OR [Abstract: "convolutional neural network"] OR [Abstract: "cnn"] OR [Abstract: "segmentation"]] |                                |   |
| <b>Refining Search</b>             | <ul style="list-style-type: none"> <li>• 2019 - 2024</li> <li>• Title and Abstract</li> </ul>  |                                |   |
| <b>Quantitative Results</b>        | 35   |                                |   |
| <b>Relevant (Title + Abstract)</b> | 24   | <b>Relevant (Full Article)</b> | 3 |

Table 4. ACM Digital Library Search.

| <b>ScienceDirect</b>               |  |                                |   |
|------------------------------------|--|--------------------------------|---|
| <b>Search Terms</b>                | ((placenta OR placental) AND (pathology) AND (deep learning OR CNN OR Transfer learning OR Segmentation) AND (Photo image analysis)) |                                |   |
| <b>Refining Search</b>             | <ul style="list-style-type: none"> <li>• 2019 - 2024</li> <li>• All Formats</li> </ul>   |                                |   |
| <b>Quantitative Results</b>        | 129  |                                |   |
| <b>Relevant (Title + Abstract)</b> | 3  | <b>Relevant (Full Article)</b> | 1 |

Table 5. ScienceDirect Search.

| <b>Pubmed</b>                      |   |                                |                    |
|------------------------------------|---|--------------------------------|--------------------|
| <b>Search Terms</b>                | (placenta) AND (pathology OR placental pathology OR gross pathology OR gross morphology) AND (deep learning OR Convolutional neural network OR Transfer learning OR Machine Learning OR Segmentation) AND (Photo image analysis)) |                                |                    |
| <b>Refining Search</b>             | <ul style="list-style-type: none"> <li>• 2019 - 2024</li> <li>• All Formats</li> </ul>  |                                |                    |
| <b>Quantitative Results</b>        | 7   |                                |                    |
| <b>Relevant (Title + Abstract)</b> | 2   | <b>Relevant (Full Article)</b> | 1 (sem duplicatas) |

Table 6. Pubmed Search.

| <b>Scielo</b>                      |   |                                |   |
|------------------------------------|---|--------------------------------|---|
| <b>Search Terms</b>                | (placenta) AND (pathology) OR (placental pathology) OR (gross pathology) OR (gross morphology) AND (deep learning) OR (transfer learning) OR (Machine Learning ) OR (CNN) OR (Segmentation) OR (Photo image analysis) |                                |   |
| <b>Refining Search</b>             | <ul style="list-style-type: none"> <li>• 2019 - 2024</li> <li>• All Formats</li> </ul>  |                                |   |
| <b>Quantitative Results</b>        | 1   |                                |   |
| <b>Relevant (Title + Abstract)</b> | 0   | <b>Relevant (Full Article)</b> | 0 |

Table 7. Scielo Search.

### 2.1.2.2 Selection Process

In the selection process of primary studies, articles published in the last six years (2019 to 2024) and written in English were considered. Additionally, the selection included relevant secondary literature, identified through the references of primary studies and citation metrics.

The selected articles must meet the quality criteria (Table 8) and inclusion criteria (Table 9) while not meeting any of the exclusion criteria (Table 10), as described below.

| Quality Criteria  |
|---|
| <ul style="list-style-type: none"><li>• Provide a detailed description of the methodology, including the architecture of deep learning models, datasets used, and evaluation metrics.</li></ul>   |
| <ul style="list-style-type: none"><li>• Include experimental validation with real data, i.e., tests conducted on physical placenta images rather than simulations or synthetic models.</li></ul>  |
| <ul style="list-style-type: none"><li>• Employ segmentation, detection, or classification techniques for placental morphological pattern recognition based on photographic images.</li></ul>      |
| <ul style="list-style-type: none"><li>• Discuss the limitations of the proposed method, including challenges faced in detecting morphological patterns.</li></ul>                                 |
| <ul style="list-style-type: none"><li>• Perform a quantitative comparison with other state-of-the-art approaches, using standardized metrics such as IoU, Dice Score, AUC-ROC, and mAP.</li></ul> |
| <ul style="list-style-type: none"><li>• Published in high-impact journals or conferences, indexed in recognized scientific databases.</li></ul>   |

Table 8. Quality Criteria.

| Inclusion Criteria   |
|--|
| <ul style="list-style-type: none"> <li>Published in English and available in full text within the selected scientific databases.</li> </ul>  |
| <ul style="list-style-type: none"> <li>Published between 2019 and 2024, ensuring that the studies reflect recent advancements in the use of deep learning for placental analysis.</li> </ul>       |
| <ul style="list-style-type: none"> <li>Application of deep neural networks (Deep Learning) for macroscopic morphological analysis of placentas using photographic images of the organ.</li> </ul>  |
| <ul style="list-style-type: none"> <li>Studies that include validation and practical experimentation, utilizing real datasets and presenting quantitative metrics for model evaluation.</li> </ul> |
| <ul style="list-style-type: none"> <li>Studies that consider clinically relevant aspects, such as accuracy in diagnosing placental alterations.</li> </ul>   |

Table 9. Inclusion Criteria.

| Exclusion Criteria   |
|--|
| <ul style="list-style-type: none"> <li>Studies that analyze macroscopic placental alterations exclusively through clinical imaging exams, such as magnetic resonance imaging (MRI), ultrasonography, or computed tomography, without the use of photographic images of the organ.</li> </ul> |
| <ul style="list-style-type: none"> <li>Studies that are purely theoretical model proposals, lacking experimental validation or demonstrated results.</li> </ul>  |
| <ul style="list-style-type: none"> <li>Articles that do not provide detailed evaluation metrics or fail to present concrete results on the performance of the models used.</li> </ul>  |

|  |
|--|
| <ul style="list-style-type: none"><li>• Studies that do not use deep learning (Deep Learning) as the primary approach for placental analysis.</li></ul>          |
| <ul style="list-style-type: none"><li>• Systematic reviews or review articles without original data or new experiments.</li></ul>                                |
| <ul style="list-style-type: none"><li>• Studies published in low-relevance scientific events, lacking peer review or recognition in indexed databases.</li></ul> |
| <ul style="list-style-type: none"><li>• Studies published before 2019.</li></ul>   |

Table 10. Exclusion Criteria.

### 3 DETAILED ANALYSIS OF THE ARTICLES

In this section, we detail the selected articles, grouped by year, with a brief description of each.

#### 3.1 2019

##### **PlacentaNet: Automatic Morphological Characterization of Placenta Photos with Deep Learning**

---

**Authors:** CHEN, Y.; WU, C.; ZHANG, Z.; GOLDSTEIN, J. A.; GERNAND, A. D.; WANG, J. Z.

**Repository:** ACM Digital Library

**Method:**

Transfer Learning (TL), Convolutional Neural Networks (CNNs), Image Segmentation, Fetal/Maternal Classification, Umbilical Cord Insertion Point Localization, Quantitative Evaluation with Standard Metrics, Comparison with Benchmark Models (U-Net, SegNet).

The study proposes a deep learning model called PlacentaNet, designed for the automated morphological analysis of placentas using photographic images. The method employs a multi-task architecture based on CNNs, with a shared encoder for three specific tasks:

- **Placental segmentation** using the SegDecoder architecture, employing Dice Loss to mitigate pixel-wise segmentation imbalances.
- **Classification of the fetal/maternal side** of the placenta through the Classification Subnet, utilizing Binary Cross-Entropy Loss (BCE).
- **Localization of the umbilical cord** insertion point using the IPDecoder, which applies a Gaussian heatmap and Mean Squared Error Loss (MSE) to predict the cord position.

The approach integrates Transfer Learning, allowing learned representations from segmentation to be reused for improving classification and localization tasks. The study validates the proposal with a dataset of 1,003 placenta images obtained from Northwestern Memorial Hospital, containing detailed annotations for segmentation and pathological analysis. Performance was evaluated using metrics such as IoU, Dice Score, BCE Loss, and PCK (Percentage of Correct Keypoints), demonstrating a significant improvement over benchmark models (U-Net, SegNet, and Hourglass).

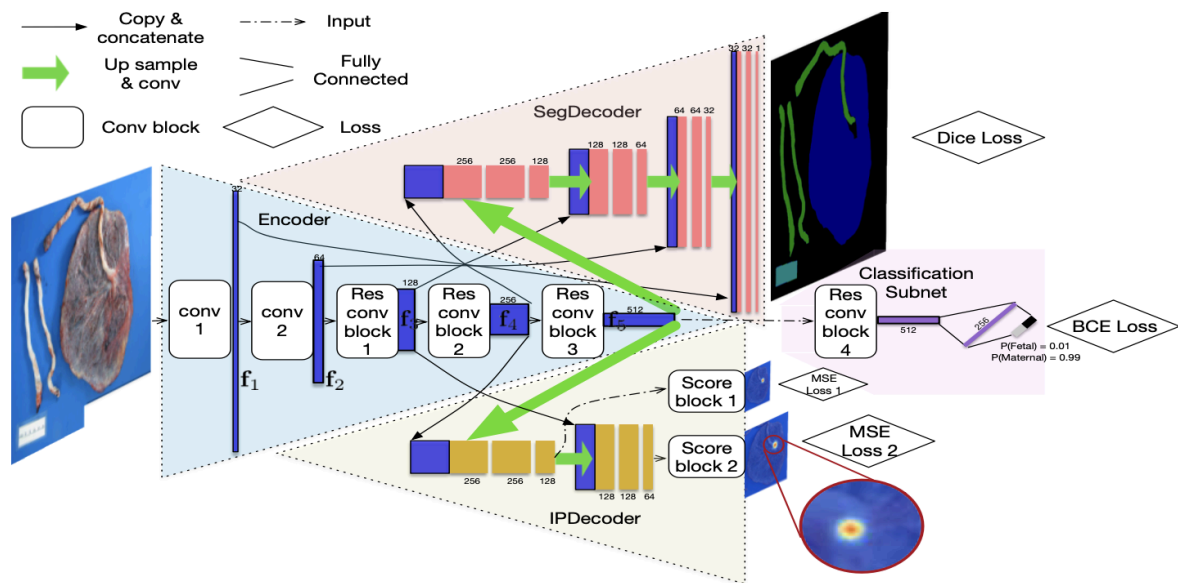


Figura 2: Arquitetura da PlacentaNet - um Codificador compartilhado entre todas as tarefas para extração de características (azul), um SegDecoder totalmente convolucional para segmentação de imagens (vermelho), uma Sub-rede de Classificação para classificação da face fetal/materna (roxo) e um IPDecoder totalmente convolucional para localização do ponto de inserção do cordão. (CHEN, Yukan et al., 2019).

### Analysis:

The PlacentaNet approach represents an innovative method for automating the morphological characterization of placentas through deep learning. The study's primary innovation lies in its multi-task architecture, which leverages shared representations across different tasks to optimize segmentation, classification, and structural localization.

#### Strengths:

- **Use of Transfer Learning:** Enhances training efficiency by enabling feature reuse from segmentation in other tasks.
- **Quantitative comparison with state-of-the-art models:** PlacentaNet outperformed U-Net, SegNet, and Hourglass in segmentation and detection metrics.
- **Real and well-documented dataset:** Images were collected from a reference hospital, ensuring practical viability for clinical applications.
- **Demonstration of clinical impact:** The study suggests applications in automated screening, potentially reducing pathologists' workload.

#### Limitations and Challenges:

- **Limited dataset size:** Despite promising results, the authors acknowledge that 1,003 images may not be sufficient to ensure model generalization across varied clinical settings.
- **Dependency on initial segmentation:** Any error in placenta segmentation may compromise subsequent classification and localization tasks.

- **Lack of validation across multiple clinical centers:** The model was tested using images from only one hospital, which may limit its applicability across different populations and imaging equipment.
- **Absence of direct clinical validation:** While results are promising, no comparative study was conducted between PlacentaNet predictions and expert pathological evaluations.

PlacentaNet emerges as a promising model for the automatic analysis of placentas, showcasing significant advancements in segmentation and morphological characterization. The study provides a solid methodological foundation for future research, particularly in the automated screening of placental anomalies. However, for clinical adoption, further studies are needed with larger and more diverse datasets, as well as the integration of clinical feedback for real-world hospital validation.

### 3.2 2020

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

**Authors:** CHEN, Y.; ZHANG, Z.; WU, C.; DAVAASUREN, D.; GOLDSTEIN, J. A.; GERNAND, A. D.; WANG, J. Z.

**Repository:** PubMed

#### Method:

The study employs complex deep learning methods, including convolutional neural networks (CNNs) and task-specific strategies for automated placental assessment and examination. A two-stage pipeline was proposed for automated placenta evaluation, as shown in Figure 3.

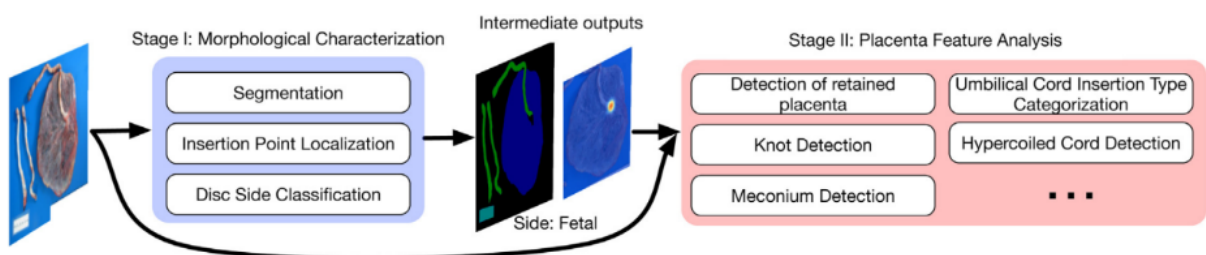


Figure 3: Schematic diagram of the proposed two-stage framework for automated placental assessment and examination using photos. (Source: CHEN, Yukun et al., 2020).

In Stage 1, for morphological characterization, the placenta was segmented using an encoder-decoder architecture, incorporating U-Net and SegNet. The proposed approach utilizes a multi-head CNN for morphological characterization tasks, employing a transfer learning strategy for training. Consequently, fetal/maternal classification and umbilical cord insertion point localization leveraged

segmentation-based information. The learned representation from segmentation improved these subsequent tasks.

In Stage 2, placental feature analysis was performed, including:

- **Retained Placenta Detection:** A classification network was used to identify retained placentas, showing promising results.
- **Umbilical Cord Insertion Type Categorization:** Classification of umbilical cord insertion type based on the distance between the insertion point and the placental margin. Achieved an overall accuracy of 88%, with precise distance estimation relative to actual values.
- **Detection of Meconium, Abruption, and Chorioamnionitis:** ROC curves were used to assess sensitivity and specificity in detecting these conditions, achieving high sensitivity and specificity.
- **Irregular Shape Detection:** Classification of placentas with irregular morphology, employing a shape irregularity metric while considering annotation subjectivity limitations.
- **Detection of Hypercoiled Umbilical Cord:** Classification of umbilical cords as hypercoiled, demonstrating reasonable sensitivity and specificity, though with opportunities for improvement.
- **Knot Detection:** Utilization of the mean average precision (mAP) metric under different overlap thresholds. Sampling strategies were explored to address class imbalances, achieving significant performance in detecting true knots.

Additionally, Stage 2 incorporated further analyses:

- **Inference Time Analysis:** Summarized inference time for each model component, highlighting system feasibility for clinical use.
- **Clinical Significance Discussion:** Explored the benefits of automated placental assessment, emphasizing objectivity, continuous availability, flexibility, and scalability in clinical applications. Identified areas for improvement and underscored the need for continuous feedback from healthcare professionals.

### **Analysis:**

The study focuses on automating placental evaluation during pregnancy, addressing the limitations of manual assessment due to time and cost constraints. It proposes a two-stage pipeline, starting with placental morphological characterization using a CNN that shares features across related tasks. The second stage focuses on detecting pathological indicators, such as retained placenta and umbilical cord insertion types. The use of transfer learning and independent models aims to enhance accessibility by leveraging photographic images. The dataset, consisting of 18,400 manually labeled images, contributes to future research. The results indicate promising efficiency gains and performance improvements, marking a significant step forward in the automation of placental assessment to benefit maternal and fetal health.

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

---

**Authors:**ZHANG, Z.; DAVAASUREN, D.; WU, C.; GOLDSTEIN, J. A.; GERNAND, A. D.; WANG, J. Z.

**Repository:** ScienceDirect

### **Method:**

Multi-Region Saliency-Aware Learning (MSL), Attention Mechanism, Saliency Constraint, Adversarial Translation, Experiments with Real-World Datasets, Efficiency Evaluation, Comparison with State-of-the-Art Approaches.

The method proposed in this study (Figure 4), referred to as MSL (Multi-Region Saliency-Aware Learning), stands out for addressing the challenge of placenta image segmentation across different domains, particularly when target domain data is limited and corresponding labels are unavailable. MSL employs a domain adaptation approach, integrating attention networks (AS and AT), generators (GS  $\rightarrow$  T and GT  $\rightarrow$  S), discriminators (DS and DT), and consistency constraints to facilitate placenta image translation between domains. A notable innovation is the introduction of attention and saliency consistency losses:

- **Attention consistency** aims to maintain **semantically crucial regions**.
- **Saliency consistency** is designed to **preserve saliency relationships** between key objects.

Additionally, adversarial losses and cycle loss are employed to enable semantic mappings. The MSL model is trained to adapt a pre-trained segmentation model to translated target domain images. The practical application of this method is highlighted by its efficacy in improving chorioamnionitis diagnosis, using pathological indicators such as fetal inflammatory response (FIR) and maternal inflammatory response (MIR).

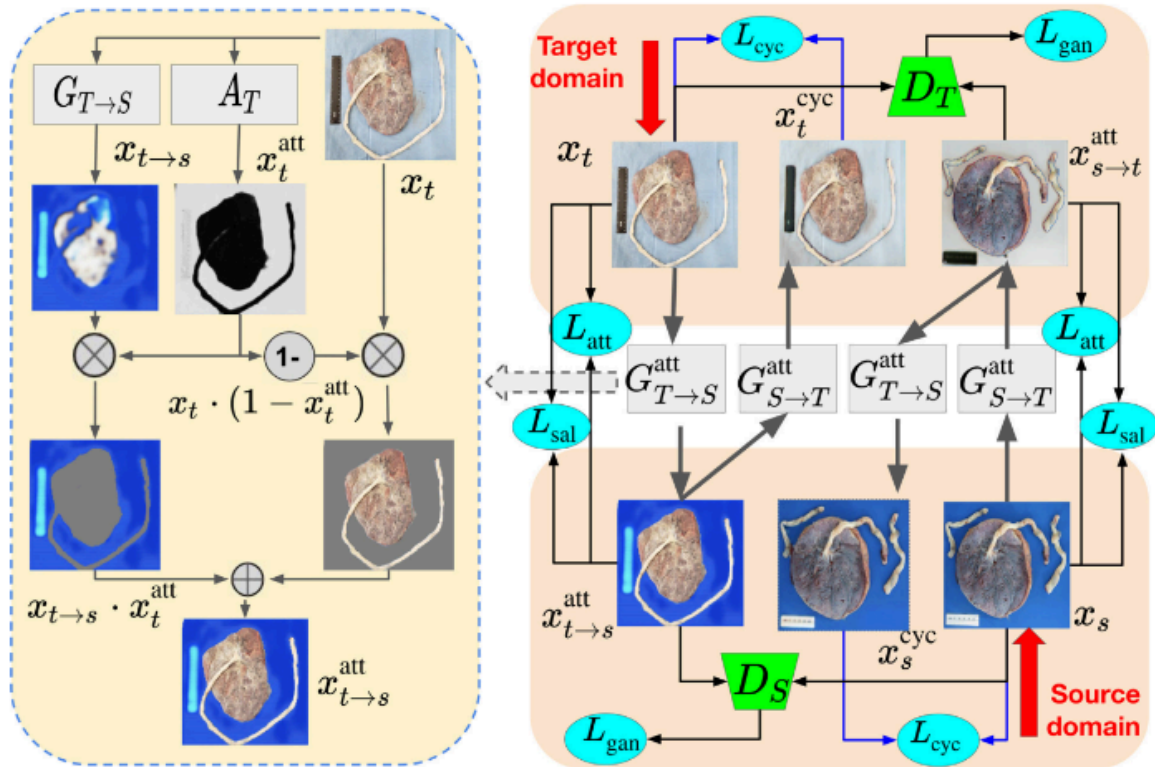


Figure 4: Proposed approach pipeline. The AT attention network divides the placenta image, highlighting regions such as the ruler and background, while preserving semantics in unattended parts. The translated image combines translated attended parts and original unattended parts, incorporating consistent attention and saliency losses, along with image-level adaptation. (Source: ZHANG, Zhuomin et al., 2020, Oct.)

### Analysis:

The innovative MSL approach integrates attention and saliency consistency constraints in domain adaptation for placenta image segmentation, making it a noteworthy contribution. These constraints proved effective in preserving crucial semantic information during cross-domain translation, demonstrating MSL's capability to handle variations in image characteristics across different hospitals. When compared to established image translation methods such as CycleGAN and AGGAN, MSL exhibited a significant improvement in segmentation accuracy, as indicated by the substantial increase in mean IoU. The quantitative analysis highlights MSL's superiority in specific segmentation scenarios, validating its effectiveness for the intended task. The practical application of this method in chorioamnionitis diagnosis enhancement further underscores its clinical relevance. The use of pathological indicators, such as FIR and MIR, reveals MSL's versatility, extending beyond image segmentation to contribute to more accurate diagnoses in various clinical settings. However, it is essential to note certain challenges that may arise, depending on data quality and availability across different clinical domains. Additionally, further research is required to generalize the method for other medical imaging modalities. Thus, MSL represents a significant contribution to domain adaptation for placenta imaging, offering an innovative and quantitatively validated approach with critical clinical implications. The inclusion of visual results, such as segmentation images and performance metrics, further reinforces confidence in the method's effectiveness.

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

---

**Authors:** PAN, Y.; GERNAND, A. D.; GOLDSTEIN, J. A.; MITHAL, L.; MWINYELLE, D.; WANG, J. Z.

**Repository:** ACM Digital Library

**Method:**

Vision-Language Contrastive Learning (VLC), NegLogCosh Similarity Metric (NegLogCosh), Sub-Feature Comparison, Pre-Training and Fine-Tuning Stages, Encoder Architectures, Experimental Evaluation.

The study proposes an advanced machine learning approach, referred to as Vision-Language Contrastive Learning (VLC), to automatically analyze photographic placenta images. This approach utilizes a contrastive training method, where the neural network learns robust placenta representations by integrating both image-based and pathological report-based information. To overcome previous challenges, the authors introduce the NegLogCosh similarity metric, which reduces the suppression of critical features. Additionally, a sub-feature comparison technique is implemented to enhance the model's sensitivity to different placental characteristics. The method is divided into two key training stages:

- **Pre-Training:** A foundational model is trained using large-scale data.
- **Fine-Tuning:** Model weights are adjusted to optimize performance on placental image analysis tasks.

Specific encoder architectures are employed, including ResNet50 for images and BERT for text-based pathology reports. The proposed approach is evaluated on various tasks, including the identification of meconium-stained placentas and inflammation patterns, demonstrating its efficacy and robustness across different acquisition conditions, including iPad-captured images.

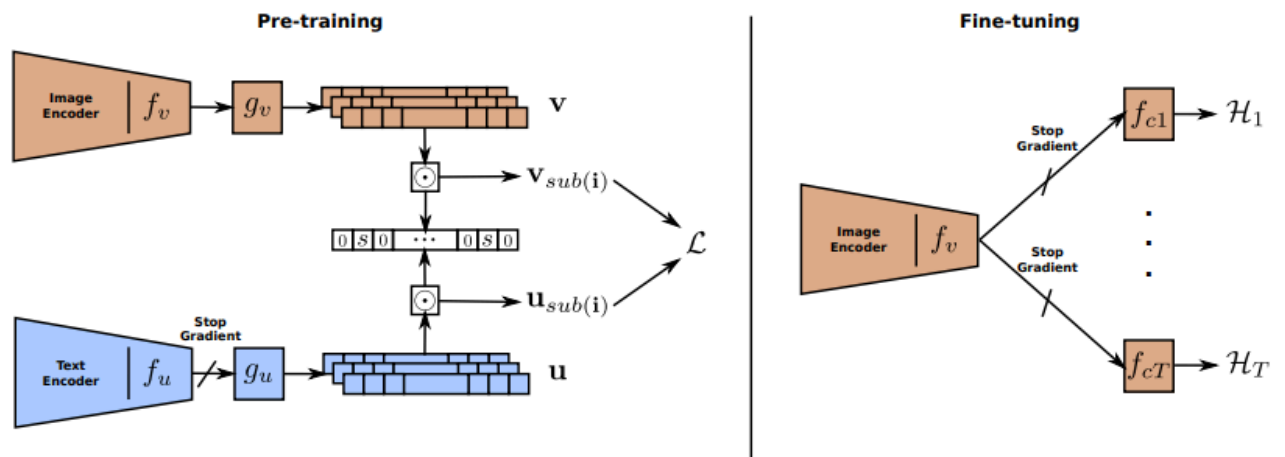


Figure 5: Proposed Method - Consists of a pre-training stage and a fine-tuning stage. The pre-trained text encoder is frozen using a gradient stop operation during the pre-training phase. The image encoder is frozen and shared across all tasks in the fine-tuning stage. (Source: PAN, Yimu et al., 2022).

### Analysis:

This study presents an innovative and advanced approach for automatic placenta analysis using machine learning techniques, particularly Vision-Language Contrastive Learning (VLC). The proposed method is designed to overcome previous challenges, such as feature suppression, with the goal of delivering more robust and efficient placenta analysis. The key innovation of this study lies in the introduction of the NegLogCosh similarity metric, which reduces the suppression of critical features during training. Additionally, the sub-feature comparison technique enhances the model's ability to distinguish between different placental characteristics. The study demonstrates promising results in placental analysis, emphasizing the importance of model generalization and robustness in diverse clinical settings. The use of specific architectures, such as ResNet50 for image analysis and BERT for text processing, further reinforces the comprehensive and specialized nature of the method. Despite these strengths, the study acknowledges the need for a larger clinical dataset for further validation and continuous improvement. Nonetheless, this work represents a significant contribution to placental analysis automation and holds great promise for future clinical applications.

3.4 2023

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

**Authors:** PAN, Y.; CAI, T.; MEHTA, M.; GERNAND, A. D.; GOLDSTEIN, J. A.; MITHAL, L.; MWINYELLE, D.; WANG, J. Z.

**Repository:** ACM Digital Library

**Method:**

Pathology Report Feature Recomposition, Vision-Language Contrastive Learning (VLC), Efficient Neural Networks, Evaluation on Different Datasets, Component Ablation, Training and Inference Efficiency.

The study proposes methods to enhance automatic placenta analysis, aiming to improve both performance and training efficiency, focusing on robust representation learning and overcoming challenges associated with the variability in pathology reports and feature distribution. An overview is presented in Figure 6.

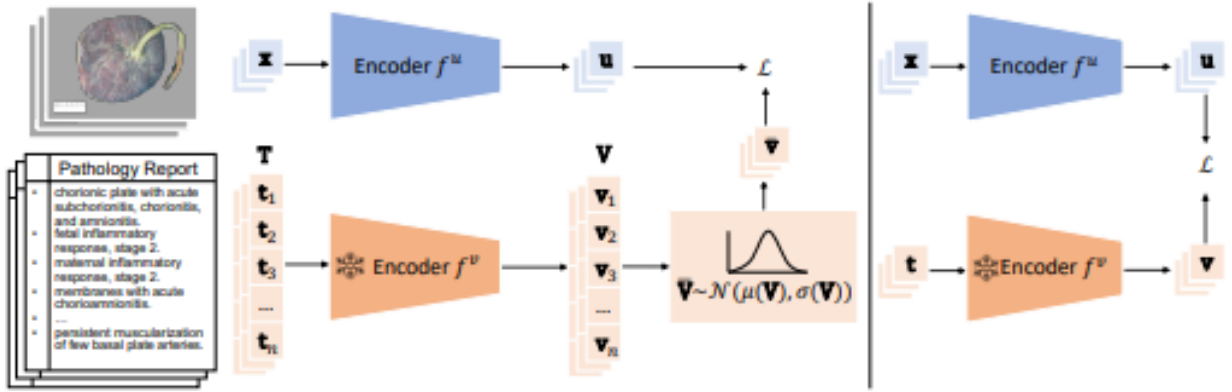


Figure 6: Comparative illustration between the proposed approach (left) and traditional Vision-Language Contrastive Learning (VLC) (right). Inputs consist of images (x) and text (t), with an example input shown on the left. (Source: PAN, Yimu et al., 2023).

The study introduces two major improvements:

1. Pathology Report Feature Recomposition
2. Distributional Feature Recomposition in Vision-Language Contrastive Learning (VLC)

The Pathology Report Feature Recomposition is a novel approach in the medical VLC domain, capturing features from pathology reports of varying lengths. This method aims to reduce feature suppression, ensuring all placental characteristics are treated equally. Additionally, the study presents a loss function formulation to achieve a more robust feature representation:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \left( \lambda \tilde{\ell}_i^{(u \rightarrow v)} + (1 - \lambda) \tilde{\ell}_i^{(v \rightarrow u)} \right)$$

The Distributional Feature Recomposition seeks to provide a more robust and distribution-aware representation. The article employs bootstrap techniques to estimate the distribution of the mean feature vector. This function is represented as  $\mathbf{N}(\mu(\mathbf{V}), \sigma(\mathbf{V}))$ , where  $\mu(\mathbf{V})$  and  $\sigma(\mathbf{V})$  denote the mean and standard deviation, respectively. During training iterations, bootstrap sampling is applied to generate a bootstrapped feature set  $\tilde{\mathbf{V}}$  from the estimated normal distribution  $\mathbf{N}(\mu(\mathbf{V}), \sigma(\mathbf{V}))$ . Thus, in

the distribution representation through the vector  $\hat{\mathbf{v}} = \sum_{\mathbf{v} \in \tilde{\mathbf{V}}} \mathbf{v}$ , the method captures the variation of features over training epochs. This approach aims to improve the representativeness of the mean vector, resulting in enhanced generalization capability. The research also leverages efficient neural networks, such as EfficientNet-B0 and MobileNetV3-Large-1.0, as backbones for the model. These lighter and faster architectures are used to improve efficiency and accessibility in placenta analysis, particularly for resource-limited healthcare settings. To assess model performance, multiple datasets were used, including an iPad-based evaluation dataset. Metrics such as AUC-ROC were calculated to measure model effectiveness for specific tasks. Additionally, training and inference efficiency were analyzed, considering factors like:

- Training time
- Throughput (examples per second)
- Floating-point operations per second (TFLOPS)

These key methodologies represent significant contributions to enhancing automated placenta analysis.

#### **Analysis:**

The article presents an innovative approach to automated placenta analysis, utilizing advanced Vision-Language Contrastive Learning (VLC) techniques. The primary contribution of the study lies in the introduction of two key improvements:

1. Pathology Report Feature Recomposition – addressing the challenge of variable-length pathology reports by reducing feature suppression.
2. Distributional Feature Recomposition – employing bootstrap techniques to stabilize feature representation across training epochs.

Regarding experimental results, the study demonstrates that the proposed approach outperforms previous methods in terms of both efficiency and performance. However, the authors highlight the need for greater robustness in evaluating iPad-acquired images, indicating potential areas for improvement. Additionally, the dependency on a large pre-trained language model may pose limitations in scalability and flexibility.

## 4 DISCUSSION

This systematic review aimed to identify, compile, and analyze academic studies that apply computer vision techniques for the identification of morphological patterns in placentas. After applying the selection criteria, five studies were included, each employing different methodologies for placental image segmentation, classification, and analysis.

The analyzed studies applied convolutional neural networks (CNNs) and various deep learning approaches to extract morphological features from placenta images. Models such as PlacentaNet (2019) and AI-PLAX (2020) utilized Transfer Learning to enhance segmentation and umbilical cord localization. In contrast, more recent studies, such as Vision-Language Contrastive Learning (2022) and Enhancing Automatic Placenta Analysis (2023), explored contrastive and multimodal learning, integrating images and clinical reports to enrich analysis capabilities.

Table 11 summarizes the technologies used in each of the analyzed studies, highlighting the methods applied for segmentation, classification, and umbilical cord localization, as well as the training strategies adopted.

| Article                             | Segmentation                | Classification                  | Umbilical Cord Localization | Training Strategy                    |
|-------------------------------------|-----------------------------|---------------------------------|-----------------------------|--------------------------------------|
| <b>PlacentaNet (2019)</b>           | SegDecoder                  | Classification Subnet           | IPDecoder                   | Transfer Learning                    |
| <b>AI-PLAX (2020)</b>               | U-Net SegNet                | Multitask CNN                   | CNN + PCK                   | CNN with Transfer Learning           |
| <b>Multi-region Saliency (2020)</b> | MSL (Multi-region Saliency) | Not addressed                   | Not addressed               | Domain Adaptation                    |
| <b>Vision-Language (2022)</b>       | ResNet50                    | BERT + Contrastive Learning     | Not addressed               | Contrastive Learning                 |
| <b>Enhancing VLC (2023)</b>         | EfficientNet-B0             | VLC + Efficient Neural Networks | Not addressed               | Contrastive Learning + Bootstrapping |

Table 11. Technologies employed in each analyzed study.

Beyond the technologies utilized, it was possible to compare the performance of models using quantitative segmentation and classification metrics. Figure 7 presents a comparative analysis of the results obtained in each study, enabling an evaluation of the effectiveness of the proposed approaches for placental analysis.

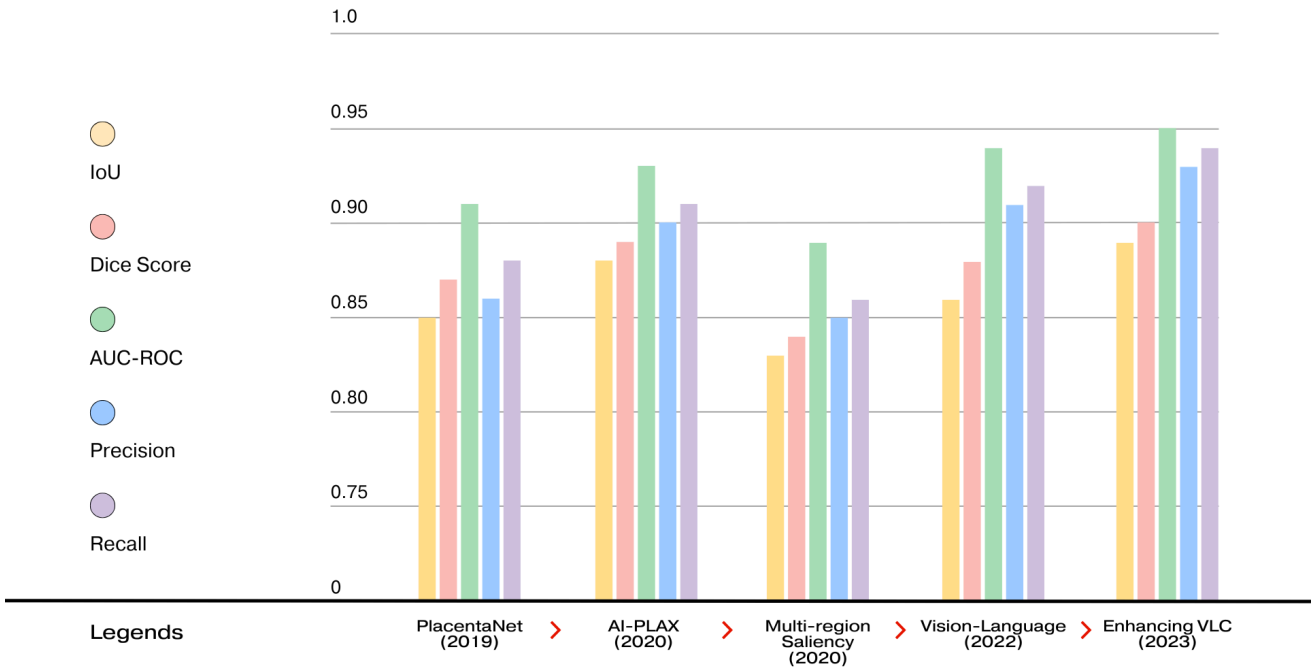


Figure 7: Comparative graph of segmentation, classification, and localization metrics for the models analyzed in this systematic review. The IoU (Segmentation), Dice Score, AUC-ROC (Classification), Precision, and Recall metrics were extracted from all studies. (Source: Authors, 2025).

In addition to the general segmentation and classification metrics, some studies also evaluated the accuracy of umbilical cord insertion point localization. Only the PlacentaNet (2019) and AI-PLAX (2020) models reported this metric, using the Percentage of Correct Keypoints (PCK) to quantify the accuracy of umbilical cord position prediction. Figure 8 presents a performance comparison of these two models in this task.

# PCK

(PERCENTAGE OF CORRECT KEYPOINTS)

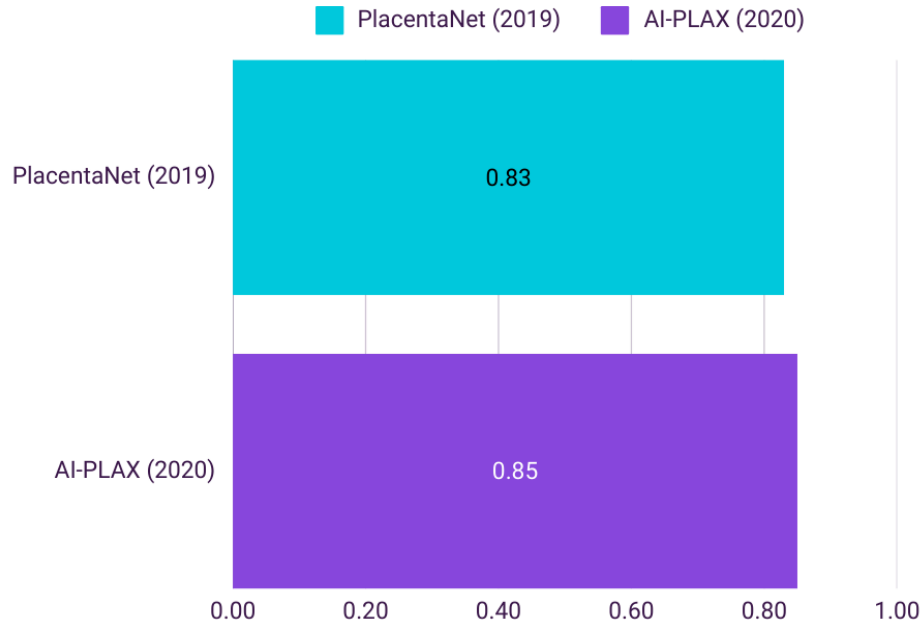


Figure 8: Comparison of the Percentage of Correct Keypoints (PCK) between PlacentaNet (2019) and AI-PLAX (2020), which performed automated detection of the umbilical cord insertion point. Other studies did not report this metric. (Source: Authors, 2025).

Based on the analyzed studies, computer vision has been shown to automatically identify morphological features relevant to placental screening and clinical analysis. The main investigated aspects include:

- **Placental disk shape** (AI-PLAX, 2020). (AI-PLAX, 2020).
- **Umbilical cord insertion patterns** (PlacentaNet, 2019).
- **Presence of meconium, placental abruption, and chorioamnionitis** (AI-PLAX, 2020).
- **Detection of retained placenta** (AI-PLAX, 2020).

From these findings, it was possible to partially answer the research questions proposed in this review:

- **What techniques are used for placental pattern recognition?**  
The analyzed studies employed CNNs (U-Net, SegNet), Transfer Learning, and Contrastive Learning (VLC) for segmentation and classification of placental images.

- **What patterns are recognized, and how are they classified?**

The research identified morphological patterns related to placental shape, structural alterations, and umbilical cord insertion.

- **What strategies are used to interpret placental shape and integrity?**

Deep learning-based segmentation, particularly using U-Net and SegDecoder, was widely applied for placental structure delineation and characterization.

- **What strategies are used for detecting placental color alterations?**

The AI-PLAX (2020) study explored ROC curves to detect chorioamnionitis and meconium through automatic color variation analysis in placenta images.

- **What is the accuracy and computational cost of these models?**

More robust models, such as AI-PLAX (2020), achieved up to 88% accuracy in anatomical pattern detection but had high computational costs, requiring specialized hardware.

- **Are these models adaptable and portable?**

PlacentaNet and AI-PLAX are highly specialized, whereas more recent approaches such as VLC (2022-2023) explore generalization across different devices and lighting conditions.

Despite the advances presented, some gaps and challenges still need to be overcome. First, the lack of publicly available databases hinders the reproducibility of research, making models highly dependent on the specific datasets used in each study. Additionally, the generalization of these models to different populations and clinical settings has not been extensively tested, as most studies rely on images from a single hospital or research center. Finally, clinical validation remains limited, as none of the analyzed studies directly compared their results with evaluations conducted by human pathologists. Therefore, although the application of computer vision in placental analysis presents advanced segmentation and classification methods, the clinical adoption of these models still requires additional studies to establish their robustness and effectiveness in real-world scenarios.

## 5 CONCLUSION

The integration of computer vision and deep learning techniques in macroscopic placenta analysis has demonstrated great potential for automating the identification of morphological patterns, increasing efficiency and accuracy in clinical assessments. This systematic review evidenced that CNN-based models, such as PlacentaNet (2019) and AI-PLAX (2020), achieved high performance in segmentation and classification, while more recent approaches, such as multimodal contrastive learning (2022-2023), explored the integration of images and clinical data to enhance interpretation and diagnostic capabilities.

The analyzed studies demonstrate that deep learning methods can segment placental structures, classify fetal and maternal surfaces, and identify key anatomical features, achieving high values in metrics such as IoU, Dice Score, and AUC-ROC. However, despite these advances, several challenges still limit the generalized clinical application of these models. Among them, the scarcity of publicly available datasets hinders research reproducibility and model scalability. Additionally, the generalization of these methods to different populations and clinical settings has not been extensively tested, as most studies rely on data from a single hospital or research center. Another limitation is the lack of large-scale clinical validation, as none of the studies directly compared their results with evaluations performed by pathologists, which is essential for ensuring reliability in medical applications.

Future research may benefit from the development of standardized datasets, validation of models in diverse clinical environments, and optimization of architectures to reduce computational costs and increase accessibility. Moreover, the incorporation of self-supervised learning and multimodal approaches has the potential to improve the robustness and interpretability of these models.

As research in this field advances, AI-assisted placental analysis may become a complementary tool to support pathologists, enabling more precise and efficient assessments and reducing the manual workload associated with anomaly screening. However, its integration into clinical practice will depend on overcoming methodological challenges and conducting rigorous large-scale validations to ensure its reliability and applicability in hospital settings.

## REFERENCES

- CHEN, Y. et al. AI-PLAX: AI-based placental assessment and examination using photos. *Computers in Medical Imaging and Graphics*, v. 84, p. 1–15, 2020.
- CHEN, Y.; WU, C.; ZHANG, Z.; GOLDSTEIN, J. A.; GERNAND, A. D.; WANG, J. Z. PlacentaNet: automatic morphological characterization of placenta photos with deep learning. In: SHEN, D. et al. (Eds.). *MICCAI 2019*. LNCS, v. 11764, p. 487–495. Springer, Cham, 2019. Disponível em: [https://doi.org/10.1007/978-3-030-32239-7\\_54](https://doi.org/10.1007/978-3-030-32239-7_54).
- HAEUSSNER, E.; SCHMITZ, C.; VON KOCH, F.; FRANK, H. G. Birth weight correlates with size but not shape of the normal human placenta. *Placenta*, v. 34, n. 7, p. 574-582, 2013. DOI: 10.1016/j.placenta.2013.04.011.
- KHONG, T. Y. et al. Sampling and definitions of placental lesions: Amsterdam placental workshop group consensus statement. *Archives of Pathology & Laboratory Medicine*, v. 140, p. 698–713, 2016.
- PAN, Y.; GERNAND, A. D.; GOLDSTEIN, J. A.; MITHAL, L. B.; MWINYELLE, D.; WANG, J. Z. Vision-Language Contrastive Learning Approach to Robust Automatic Placenta Analysis Using Photographic Images. *MICCAI (3)*, p. 707-716, 2022.
- PAN, Y.; CAI, T.; MEHTA, M.; GERNAND, A. D.; GOLDSTEIN, J. A.; MITHAL, L. B.; MWINYELLE, D.; GALLAGHER, K.; WANG, J. Z. Enhancing Automatic Placenta Analysis Through Distributional Feature Recomposition in Vision-Language Contrastive Learning. *MICCAI (6)*, p. 116-126, 2023.
- PTACEK, I.; SEBIRE, N. J.; MAN, J. A.; BROWNBILL, P.; HEAZELL, A. E. Systematic review of placental pathology reported in association with stillbirth. *Placenta*, v. 35, n. 8, p. 552–562, 2014.
- ROBERTS, D. J. Placental pathology, a survival guide. *Archives of Pathology & Laboratory Medicine*, v. 132, p. 641–651, 2008.
- SALAFIA, C. M. et al. Placental surface shape, function, and effects of maternal and fetal vascular pathology. *Placenta*, v. 31, n. 11, p. 958-962, 2010. DOI: 10.1016/j.placenta.2010.09.005.
- SILVA, T. et al. A importância do exame anatomopatológico da placenta em obstetria e neonatologia. *HU Revista*, Juiz de Fora, v. 42, n. 3, p. 171-175, 2016.
- YAMPOLSKY, M.; SALAFIA, C. M.; SHLAKHTER, O.; HAAS, D.; EUCKER, B.; THORP, J. Centrality of the umbilical cord insertion in a human placenta influences the placental efficiency. *Placenta*, v. 30, n. 12, p. 1058-1064, 2009. DOI: 10.1016/j.placenta.2009.10.001.

YAMPOLSKY, M.; SALAFIA, C. M.; SHLAKHTER, O.; HAAS, D.; EUCKER, B.; THORP, J. Modeling the variability of shapes of a human placenta. *Placenta*, v. 29, n. 9, p. 790-797, 2008. DOI: 10.1016/j.placenta.2008.06.005.

ZHANG, Z.; DAVAASUREN, D.; WU, C.; GOLDSTEIN, J. A.; GERNAND, A. D.; WANG, J. Z. Multi-region saliency-aware learning for cross-domain placenta image segmentation. *Pattern Recognition Letters*, v. 140, p. 165–171, 2020.

**APÊNDICE B – ADVANCING AUTOMATED PLACENTAL SCREENING: DEEP LEARNING FOR MULTICLASS SEGMENTATION IN POSTPARTUM IMAGE**

**Autores:** Beatriz Silva Lopes, Bibiana Quatrin Tiellet da Silva, Aldo von Wange-  
nheim, Stephan Krug e Alex Sandro Roschildt Pinto.

**Aceito e Apresentado:** 2025 LI Centro Latinoamericano de Estudios en Infor-  
mática (CLEI).

# Advancing Automated Placental Screening: Deep Learning for Multiclass Segmentation in Postpartum Images

1<sup>st</sup> Beatriz Silva Lopes  
*dept. Informática e Estatística (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brazil  
0009-0005-7700-1719

2<sup>nd</sup> Bibiana Quatrin Tiellet da Silva  
*dept. Informática e Estatística (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brazil  
0000-0001-8774-5643

3<sup>rd</sup> Aldo von Wangenheim  
*dept. Informática e Estatística (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brazil  
0000-0003-4532-1417

4<sup>th</sup> Stephan Krug  
*dept. Informática e Estatística (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brazil  
stephan.krug@grad.ufsc.br

5<sup>th</sup> Alex Sandro Roschildt Pinto  
*dept. Informática e Estatística (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brazil  
a.r.pinto@ufsc.br

**Abstract**—Postpartum placental assessment is essential for clarifying adverse pregnancy outcomes and informing clinical decisions; however, anatomopathological examination is generally reserved for selected cases due to structural and operational constraints. This study presents a deep learning-based pipeline for segmenting placental structures in real images taken in the delivery room, aiming to support clinical triage. A proprietary dataset was built using a standardized photographic protocol and annotated by pathologists across nine morphological classes. Five architectures were evaluated: U-Net with ResNet34, ResNet50, EfficientNet-B0, and EfficientNet-B7 backbones, in addition to YOLOv11 for instance segmentation. ResNet34 achieved the best overall performance (Dice: 81.2%, IoU: 70.6%, Accuracy: 85.3%), while YOLOv11 reached a AP50 of 73.2% in detecting key anatomical components. Despite the limitation imposed by the small dataset—which may affect the models’ generalization capability—the results demonstrate the feasibility of using AI for photographic placental triage, with potential to assist clinical decisions and optimize resource use in obstetric settings.

**Index Terms**—Placenta, Deep Learning, Pathology, Instance Segmentation, Postpartum Screening.

## I. INTRODUCTION

The placenta is an essential organ for fetal development, performing critical metabolic, immunological, and hormonal functions. Morphological alterations are associated with various obstetric complications, such as intrauterine growth restriction, preeclampsia, and fetal death [1]–[4]. Histopathological examination is a valuable tool for identifying these alterations and clarifying adverse perinatal outcomes [5], [6].

However, the systematic analysis of all placentas through histopathological examination is unfeasible in most hospitals due to structural limitations, shortages of human resources, and

operational costs. As a result, such examinations are typically reserved for cases with defined clinical criteria or evident visual abnormalities at delivery. In this context, artificial intelligence (AI), particularly deep learning techniques, emerges as a promising strategy to automate initial screening and support decisions regarding the need for further testing.

Although AI has been successfully applied to medical imaging, its use in placental pathology has mostly focused on prenatal examinations (such as ultrasound and magnetic resonance imaging) [7]–[10], or controlled laboratory settings [11]–[15]. The lack of solutions targeting postnatal macroscopic analysis was confirmed by a systematic review conducted by our team [17], which highlighted a gap between computational advances and their practical implementation in hospital environments.

To address this gap, we proposed a standardized photographic protocol for placental image acquisition in the delivery room [18], considering parameters such as background, lighting, and positioning. This protocol has been employed to construct a proprietary dataset consisting of real images annotated by pathologists into relevant morphological classes.

This study builds upon our previously accepted work presented at the IEEE International Symposium on Computer-Based Medical Systems (CBMS) 2025 [31], where three convolutional network architectures were evaluated for placental segmentation in a clinical setting. In the present article, we extend this research with new experiments, including hyperparameter tuning, evaluation at higher resolutions, and the addition of new architectures—such as U-Net with ResNet50 and EfficientNet-B7—without direct comparison to previous results.

Our goal is to demonstrate technical and methodological advancements toward a robust tool suitable for obstetric rou-

This work was performed with computational resources provided by FAPESC - Foundation for Research and Innovation Support of the State of Santa Catarina, grant nr. 2024TR000090

## MACROSCOPIC ANALYSIS OF THE PLACENTA AND ANATOMOPATHOLOGICAL EXAMINATION

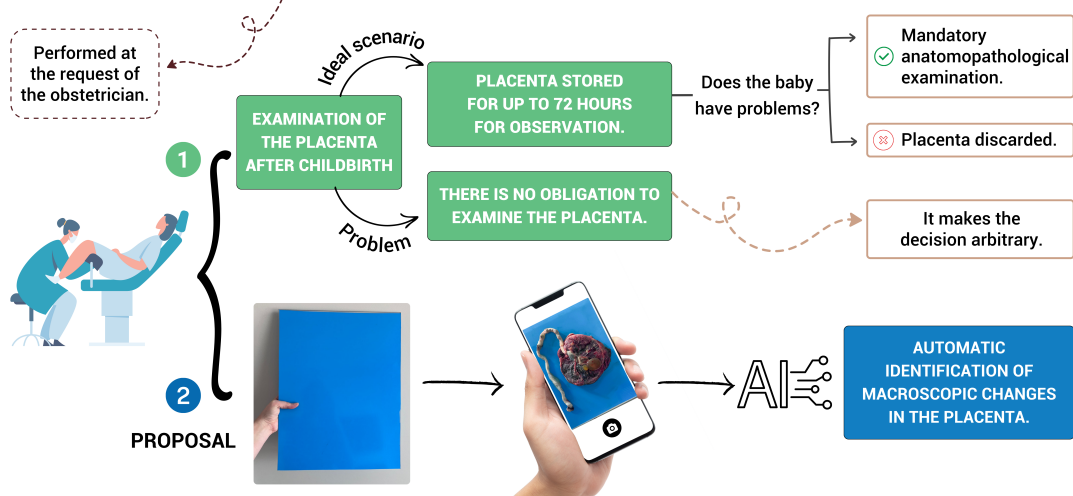


Fig. 1. Overview of the current workflow for macroscopic placental assessment and the proposed AI-based solution. After childbirth, the birth attendant (e.g., physician, midwife, or nurse) performs an initial morphological examination to identify gross abnormalities or clinical issues in the newborn or mother. If alterations are found, the placenta is sent for anatomopathological examination (AP). If no alterations are observed and there is no adverse clinical evolution, the placenta is stored for a short period (typically 72 hours to one week) and then discarded. The proposed approach introduces an automated screening pipeline, using standardized photographic capture at the time of delivery and AI-based identification of relevant morphological features to support clinical decision-making.

time, reinforcing the role of AI as a clinical decision support system at the point of care.

### A. Related Work

The application of artificial intelligence (AI) to placental analysis has predominantly focused on prenatal imaging modalities, such as ultrasound and magnetic resonance imaging, typically conducted in controlled environments [1]–[3]. Few studies employ post-delivery photographic images, and those that do generally address specific tasks such as cord insertion classification or placental maturity assessment [11], [12].

Models such as PlacentaNet [11] and AI-PLAX [12] propose multi-task pipelines but face limitations related to image standardization, clinical validation, and the scarcity of publicly annotated datasets [3], [8]. More recent strategies, including those by Pan et al. [14], [15], explore contrastive learning approaches combining images and clinical text to enhance robustness across diverse hospital contexts.

However, studies applying modern segmentation techniques to real-world macroscopic placental images remain scarce. Some approaches, such as that of Zhang et al. [13], leverage domain adaptation and saliency-aware learning, but their focus is limited to specific placental regions.

This work aims to advance the field by evaluating several state-of-the-art segmentation and detection architectures—including U-Net with ResNet34 and ResNet50 backbones, EfficientNet-B0 and B7, and YOLOv11 [24], [25]—on clinical images captured using a standardized photographic protocol [18]. The models were assessed using robust metrics, including Dice [28], IoU, accuracy, training/validation loss,

and confusion matrix, enabling a comprehensive evaluation across multiple morphological structures.

### B. Contributions

The main contributions of this work can be summarized as follows:

- **Extension of a previously validated pipeline with new architectures:** A previously validated pipeline [31] was expanded by incorporating modern architectures, such as EfficientNet-B7 and ResNet50, alongside hyperparameter adjustments including batch size and number of epochs;
- **Re-evaluation of models using refined data:** Previously evaluated networks (ResNet34 and EfficientNet-B0) were retrained with improved preprocessing and annotation procedures, resulting in enhanced performance compared to earlier results;
- **Enhancement of the clinical dataset:** Although the image set remained the same, updates in annotations and the data loading structure contributed to more consistent outcomes, emphasizing the importance of a well-defined pipeline;
- **Standardized comparative evaluation by architecture type:** Segmentation models (ResNet34, ResNet50, EfficientNet-B0, and B7) were compared under a unified experimental protocol using consistent metrics, such as Dice coefficient (the primary criterion for best model selection), IoU, accuracy, and training/validation loss. The detection model (YOLOv11) was evaluated using task-specific metrics, including AP50, AP50–95, confusion matrix, and visual predictions;

- **Advancement in AI applications for maternal health:**

This study advances an ongoing research line dedicated to automated postnatal placental screening, with an emphasis on low-cost solutions compatible with point-of-care settings.

## II. MATERIALS AND METHODS

This section describes the dataset, image acquisition protocol, annotation process, network architectures, training configurations, and evaluation metrics.

### A. Dataset and Annotation Process

This study employed a dataset composed of 156 original images captured immediately after delivery at partner obstetric centers and maternity hospitals. The images were collected under two distinct conditions: one subset was acquired under routine hospital procedures, while the other followed a standardized protocol developed by our team, which involved a neutral background, diffuse lighting, and orthogonal framing [18].

Images were organized and annotated using the Roboflow platform [32], which provides interactive tools for annotation, dataset management, image transformation, and export in various formats. Annotations targeted the main clinically relevant morphological structures of the placenta. However, only the classes with a sufficient number of representative visual examples were used in the experiments. Table 1 presents the annotation distribution across the morphological classes.

| CLASS NAME               | COUNT |
|--------------------------|-------|
| cordao_umbilical_fetal   | 83    |
| cordao_umbilical_materno | 72    |
| disco_discoide_fetal     | 56    |
| disco_discoide_materno   | 29    |
| disco_irregular_fetal    | 7     |
| disco_irregular_materno  | 8     |
| disco_ovalada_fetal      | 16    |
| disco_ovalada_materno    | 41    |
| vasos_arboriforme        | 515   |

TABLE I  
DISTRIBUTION OF LABELED INSTANCES ACROSS MORPHOLOGICAL CLASSES USED FOR TRAINING.

For the semantic segmentation experiments—primarily based on the U-Net architecture—the dataset was expanded from 156 to 468 images using data augmentation techniques [19]. The applied transformations included rotations ( $\pm 15^\circ$ ), horizontal and vertical flips, shear distortion ( $\pm 10\%$ ), saturation variations ( $\pm 25\%$ ), and exposure adjustments ( $\pm 10\%$ ). These transformations are exemplified in Figure 2.

During preprocessing, a misalignment was identified between some RGB images and their corresponding masks,

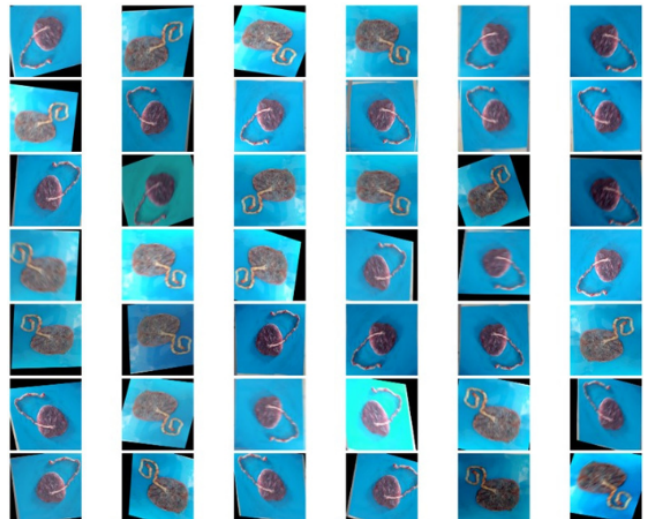


Fig. 2. Examples of synthetic images generated through data augmentation, simulating common variations in mobile device captures (rotation, lighting, geometric distortion, and saturation).

caused by the *EXIF Orientation* metadata. This metadata preserves the original capture orientation (portrait or landscape) without physically rearranging the pixel matrix on disk. While this behavior is common in digital cameras to speed up image encoding, it may lead to inconsistencies when computer vision applications—or image-reading libraries—fail to interpret the orientation field correctly. In practice, this results in images and masks with swapped dimensions or axes (e.g., pixels stored as  $x, y$  being displayed as  $y, x$ ), thus feeding the model with incorrectly aligned data.

To correct this and ensure consistency between each image and its mask, the *Auto-Orient* feature from Roboflow was applied during preprocessing. This function removes rotation metadata and rewrites the pixel matrix in the correct final orientation, standardizing all images regardless of the camera position at capture time [33]. After normalization, images were resized to  $1600 \times 1200$  px and converted to lossless PNG format, ensuring the preservation of relevant morphological details.

This adjustment was applied only in the semantic segmentation experiments (class-specific masks). In the instance detection task using the YOLOv11 architecture [24], the annotations consisted of bounding boxes independent of the masks, and therefore did not present the same issue. The data were exported in the following formats:

- **MASKS:** for semantic segmentation models, with one binary mask per class;
- **YOLO11 (TXT):** for YOLOv11, following the official architecture standard with normalized class and position coordinates [25].

### B. Photographic Capture Protocol

To standardize the images and ensure their suitability for artificial intelligence models, a specific protocol was developed

for photographing placentas immediately after delivery, in accordance with ethical principles approved by the Research Ethics Committee (CEP/CONEP).

Images were captured before the placenta was sent to the pathology department or discarded. The organ was cleaned and placed on a transparent acrylic plate with a blue vinyl background, providing suitable contrast and visual isolation. Lighting was provided by ambient or directed artificial sources, avoiding shadows and reflections.

Photographs were taken using Android or iOS mobile devices with rear cameras of at least 12 MP resolution. The device was held perpendicular to the plate to capture the entire organ area, ensuring standardized distance and framing across different records.

For each placenta, images were acquired from both main sides: the fetal side (showing vessels, membranes, and the umbilical cord) and the maternal side (displaying cotyledons and septa), based on anatomical landmarks. When possible, hands, instruments, or other objects were kept out of the frame.

This protocol proved effective for building a homogeneous image dataset that meets the technical requirements for segmentation and automated classification algorithms [18]. An example of standardized versus unstandardized capture is shown in Figure 3.

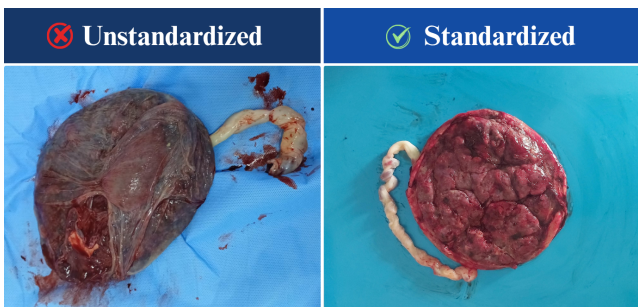


Fig. 3. Comparison between a unstandardized image (captured without adherence to the photographic protocol) and a standardized image (captured in accordance with the protocol's recommended procedures).

### C. Architectures Used

This study evaluated five deep learning architectures applied to semantic segmentation and instance detection in macroscopic placental images. The selected networks are well-established in the medical imaging domain and exhibit complementary characteristics in terms of depth, computational efficiency, and representational capacity.

**1) U-Net with ResNet34 and ResNet50:** U-Net is a widely adopted architecture for biomedical segmentation tasks, combining an encoder-decoder structure with skip connections to preserve spatial details [22]. In this study, we used U-Net variants with ResNet34 and ResNet50 backbones. ResNet34 is known for its residual connections, which mitigate vanishing gradients in deep networks, while ResNet50 introduces bottleneck blocks, allowing for increased depth with reduced computational cost [3], [4], [26]. The use of both variants

aimed to explore the trade-off between learning capacity and stability when processing clinical images with noise and morphological variability.

**2) U-Net with EfficientNet-B0 and EfficientNet-B7:** EfficientNet variants use compound scaling to balance network depth, width, and resolution, achieving high performance with fewer parameters [27]. EfficientNet-B0—the baseline model of the family—was re-evaluated in this study using refined annotations. EfficientNet-B7, a deeper and more expressive variant, was tested for its potential advantages in morphological placental screening.

**3) YOLOv11:** was used for the instance detection task. This single-stage architecture simultaneously predicts classes and bounding boxes, offering high speed and good accuracy—features desirable in clinical operational settings [24], [25]. As it differs from U-Net-based segmentation models, its evaluation included metrics such as AP50, AP50–95, confusion matrix, and qualitative prediction outputs.

### D. Training Configurations

All experiments were conducted on an NVIDIA DGX workstation equipped with an H100 GPU, enabling efficient training of deep networks using high-resolution images. For each architecture, training configurations were tailored to its complexity, aiming to balance performance, computational cost, and generalization capability.

The environment was developed in Python, using the PyTorch framework as the foundation for model construction and training [29]. The main libraries used are listed below:

- **FastAI:** a high-level API built on PyTorch, employed in the design and training of U-Net models with ResNet and EfficientNet backbones [2];
- **Albumentations:** used for data augmentation, providing fast and optimized transformations for computer vision tasks [19];
- **LAPIXDL:** a Brazilian library for visualization and evaluation of medical segmentations [35];
- **Ultralytics:** the official package for training YOLOv11 models [25];
- **Weights & Biases (W&B):** a platform used for tracking and monitoring training metrics [36];
- **EfficientUNet-PyTorch:** a library used to integrate EfficientNet backbones with U-Net [37];
- **Matplotlib, Pandas, and Seaborn:** used for exploratory data analysis, plotting, and confusion matrix visualization [8]–[10].

The U-Net models with ResNet and EfficientNet backbones were trained in three successive stages, with progressively increasing image resolutions ( $\frac{1}{4}$ ,  $\frac{1}{2}$ , and full), combined with fine-tuning. The number of epochs, batch size, loss function, and optimizer were adjusted based on the observed performance at each stage.

Table 2 summarizes the main parameters adopted in each experimental configuration, including the architecture, training strategy, input resolution, and dataset split.

|              | Architecture / Pretrained Model | Backbone        | Training strategy                       | Input image size | Batch Size | Epochs              | Optimizer | Number of images (Train/val/test) |
|--------------|---------------------------------|-----------------|---|------------------|------------|---------------------|-----------|-----------------------------------|
| Yolo         | Yolov11                         | YOLO11n-seg     | Continuous training                     | 640x640          | 64         | 200                 | AdamW     | 330/31/16                         |
| ResNet       | U-Net                           | ResNet34        | 3-stage transfer learning + fine-tuning | 1600x1200        | 8/4/2      | 50 + 50 fine-tuning | AdamW     | 376/46/46                         |
|              |                                 | ResNet50        |   |                  |            | 30 + 15 fine-tuning |           |                                   |
| EfficientNet | U-Net                           | EfficientNet-B0 | 3-stage Multiresolution + fine-tuning   | 512x512          | 16         | 50 + 50 fine-tuning | AdamW     | 374/47/47                         |
|              |                                 | EfficientNet-B7 |   |                  |            |                     |           |                                   |

TABLE II  
TRAINING SETTINGS PER MODEL

### E. Evaluation Metrics

The performance of the models was assessed using widely recognized metrics for segmentation and detection tasks, aiming to capture both overall quality and per-class behavior.

For semantic segmentation models (U-Net with different backbones), the following metrics were adopted:

- **Dice Coefficient (F1-Score):** Measures the similarity between predicted and ground truth masks; particularly sensitive to minority classes [21];
- **Intersection over Union (IoU):** Also known as the Jaccard Index; evaluates the ratio of intersection over union between segmented regions [2];
- **Foreground Accuracy:** Calculates accuracy excluding the background class, highlighting performance on clinically relevant morphological structures [3].

For the instance segmentation model (YOLOv11), the following were used:

- **AP50:** Average precision considering an IoU threshold  $\geq 0.5$ ;
- **AP50–95:** Average precision across multiple thresholds (from 0.5 to 0.95 in increments of 0.05), providing a more stringent evaluation [28].

Additionally, the following analyses were conducted:

- **Loss curves (train/val):** Tracked across epochs to monitor learning dynamics, training stability, and signs of overfitting;
- **Multiclass confusion matrices:** Used to analyze frequent misclassifications between visually similar classes;
- **Qualitative visualizations:** Compared ground truth and predicted masks to support subjective performance assessment.

The combination of objective metrics and visual analyses enabled a comprehensive evaluation of model performance, allowing for a thorough comparison between different architectures.

### III. RESULTS

Five architectures were evaluated with a focus on segmenting morphological structures of the placenta: U-Net with ResNet34, ResNet50, EfficientNet-B0, EfficientNet-B7, and YOLOv11. The U-Net-based networks followed a three-stage training strategy with increasing image resolutions ( $\frac{1}{4}$ ,  $\frac{1}{2}$ , and full), which allowed for improved spatial refinement of the predicted masks [18].

Table 3 summarizes the main performance metrics—Dice, IoU, and accuracy—obtained for each U-Net architecture in the placental segmentation tasks.

|                 | Dice   | IoU    | Accuracy |
|-----------------|--------|--------|----------|
| ResNet34        | 81,19% | 70,58% | 85,31%   |
| ResNet50        | 61,35% | 50,51% | 73,01%   |
| EfficientNet-B0 | 74,55% | 68,15% | 89,77%   |
| EfficientNet-B7 | 75,29% | 69,09% | 90,40%   |

TABLE III  
SUMMARY OF PERFORMANCE METRICS FOR U-NET MODELS IN PLACENTAL SEGMENTATION

For the instance segmentation task, YOLOv11 achieved a AP50 of 73.2% and a AP50–95 of 60.2%, with higher precision in classes such as the umbilical cord and oval discs [24], [25]. In contrast, the arborized vessels class exhibited lower accuracy, reflecting its greater morphological complexity and a possible class imbalance in the dataset. Figure 4 presents a summary of these metrics by class, highlighting the performance variations among the different morphological structures analyzed.

Figure 5 shows segmentation examples generated by ResNet34 and EfficientNet-B7, demonstrating good delineation of the main anatomical structures.

Figure 6 presents the confusion matrix and loss curves for the ResNet34 model, serving as an example to illustrate class-wise performance and training progression over the epochs.

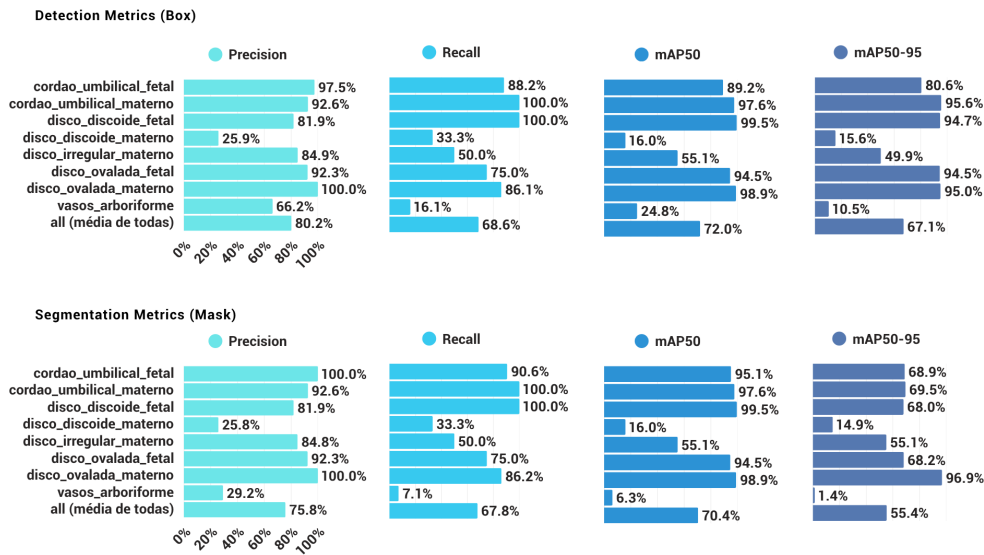


Fig. 4. Detection (Box) and segmentation (Mask) metrics obtained by the YOLOv11 architecture, presented by class. The metrics include Precision, Recall, mAP@50, and mAP@50-95, showing higher performance in structures such as the umbilical cord and oval discs, and lower accuracy in arborized vessels.

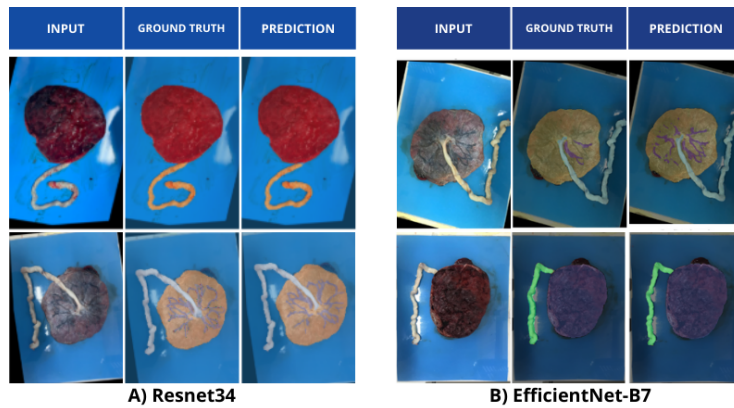


Fig. 5. Segmentation examples produced by ResNet34 and EfficientNet-B7. Each row shows, from left to right: input image, ground truth, and predicted mask.

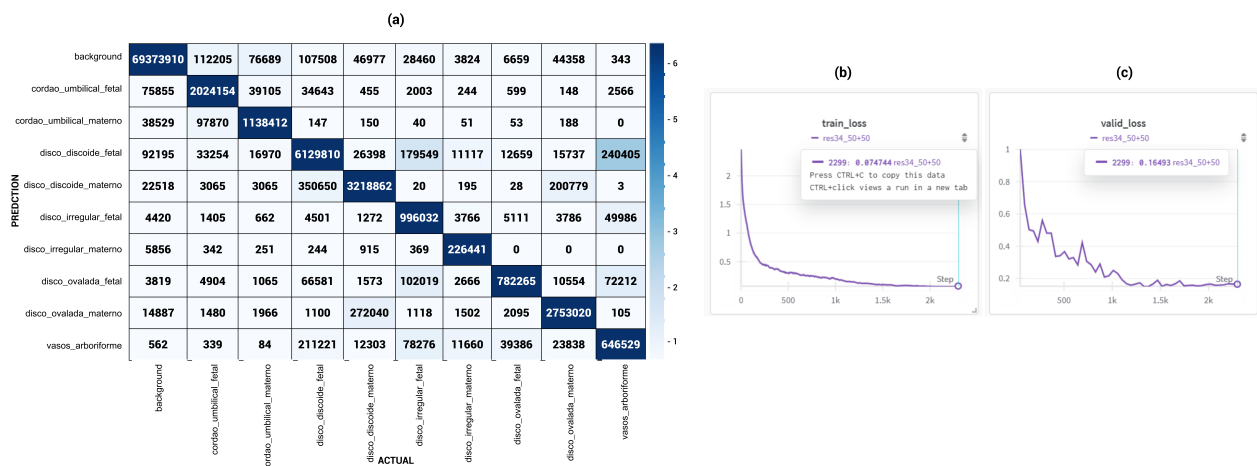


Fig. 6. Evaluation example for the U-Net model with a ResNet34 backbone. (a) Multiclass confusion matrix on the test set, showing patterns of correct predictions and inter-class confusion. (b) Training loss curve. (c) Validation loss curve.

#### IV. DISCUSSION

Unlike most previous studies that rely on ultrasound, magnetic resonance imaging, or controlled laboratory environments [7]–[9], this work proposes the application of artificial intelligence techniques to real photographs captured in the delivery room. This approach aims to reflect real-world clinical settings, thereby enhancing practical applicability. However, it also introduces significant challenges to segmentation, such as irregular lighting, the presence of blood, and visually similar structures. To mitigate these factors, a standardized photographic protocol was adopted [18], which had a direct impact on the visibility of morphological structures, facilitating both annotation and automated learning—particularly in the detection of vessels and umbilical cords, which are critical due to their high anatomical variability.

In terms of model performance, U-Net with a ResNet34 backbone achieved the best balance among accuracy, stability, and computational cost—even when applied to a dataset affected by noise and class imbalance—confirming findings from previous biomedical segmentation literature [38], [39]. ResNet50, in contrast, did not yield significant improvements over its shallower counterpart, suggesting that deeper architectures do not necessarily translate into better results in clinical scenarios with limited data volume. The EfficientNet-B0 and B7 variants, however, demonstrated solid performance after targeted hyperparameter adjustments—such as using a batch size of 16—indicating that model success depends not only on the architecture itself but also on fine-tuning and training strategies.

The YOLOv11 architecture proved effective for instance segmentation tasks, achieving a AP50 of 73.2% and AP50–95 of 60.2%, with stronger performance in well-defined structures such as the umbilical cord and placental disc. Nonetheless, the network exhibited limitations in classes such as arborized vessels, suggesting that morphological complexity and class imbalance negatively affect the model’s ability to learn consistent patterns. Complementary analyses using confusion matrices and qualitative visualizations were essential to understand these behaviors, reinforcing the importance of per-class evaluation—rather than relying solely on global metrics—in biomedical models [28], [31].

As highlighted in our systematic literature review [17], there remains a lack of studies exploring multiclass segmentation in real macroscopic placental images. This work contributes to addressing that gap by demonstrating that multiresolution approaches, standardized acquisition protocols, multicriteria evaluation, and progressive training are key components for advancing practical AI-based solutions in real clinical contexts.

From a clinical perspective, healthcare professionals involved in childbirth care—including physicians, obstetric nurses, and doulas—are naturally interested partners in the application of tools that can support rapid and accurate decision-making [41]–[43]. Adverse obstetric outcomes have a critical impact, as it is generally not expected that something will

deviate from the normal course of childbirth [44], [45]. At the same time, the high demand for care and the limitation of resources in healthcare services reinforce the utility potential of technological solutions capable of optimizing case screening and prioritization [46]. Although this study did not include field validation, integrating the model into real clinical workflows, along with collecting feedback from end users, will be essential to assess its impact on decision-making and its feasibility for large-scale adoption (AI-plax and PlacentaNet).

##### A. Threats to Study Validity

The main threat to the validity of this study concerns the reduced size and clinical limitation of the dataset used. Although the images represent real delivery room conditions, the representation of lesions associated with a wider range of clinical contexts is still limited, restricting the variety of possible diseases that can be recognized by the models.

The placenta is a morphologically complex organ, whose analysis involves recognizing a wide range of structural patterns and variations within normality. Relevant alterations may manifest in subtle and distinct ways, requiring a dataset that captures not only different pathological presentations but also the inherent anatomical variability of the organ.

The reduced size and imbalance among anatomical classes in the dataset limit the models’ generalization capacity, especially for rare morphological anomalies. In this study, such imbalance may have affected the learning of certain structures, such as the arborized vessels class. The next step should focus on expanding the dataset to incorporate greater diversity of cases and clinical contexts, thereby improving segmentation fidelity and enhancing the robustness of the models, including field validation for real clinical application.

#### V. CONCLUSION AND FUTURE WORK

This study explored the application of deep learning architectures for multiclass segmentation in macroscopic images of the placenta captured directly in the delivery room. Among the evaluated models, U-Net integrated with ResNet34 achieved the best overall performance, while EfficientNet-based variants demonstrated solid stability following hyperparameter adjustments.

Despite these encouraging results, the study was limited by a small and imbalanced dataset and the inherent challenge of segmenting morphologically similar structures. Class-wise performance analyses using confusion matrices reinforced these limitations, highlighting the importance of targeted evaluation metrics in biomedical modeling.

Future work will focus on:

- **Expanding the dataset**, focusing on greater representativeness and balance among classes, including diversity of clinical contexts and capture conditions (both protocolized and non-protocolized), in order to increase the robustness and generalization capacity of the models;
- **Enhancing annotations** with greater structural detail, especially in morphologically ambiguous regions;

- **Conducting in-depth assessments** of HRNet and exploring attention-based architectures such as transformers;
- **Pursuing clinical validation**, with the goal of integrating these models into a decision-support system—potentially via a mobile application—to assist in the automated triage of placentas at the point of care.

#### ACKNOWLEDGMENT

We would like to thank the healthcare professionals who supported the image capture process and contributed to the construction of our dataset.

We are especially grateful to all the postpartum women who consented to the photographic documentation of their placentas, enabling this research.

This work was performed with computational resources provided by FAPESC — Foundation for Research and Innovation Support of the State of Santa Catarina (grant nr. 2024TR000090).

We also acknowledge CNPq — the Brazilian National Council for Scientific and Technological Development — for the graduate scholarship granted to support this research.

#### REFERENCES

- [1] D. J. Roberts, "Placental pathology, a survival guide," *Archives of Pathology & Laboratory Medicine*, vol. 132, no. 4, p. 641–651, Apr. 2008. [Online]. Available: <http://dx.doi.org/10.5858/2008-132-641-PPASG>
- [2] V. Costanzo, A. Bardelli, S. Siena, and S. Abrignani, "Exploring the links between cancer and placenta development," *Open Biology*, vol. 8, no. 6, Jun. 2018. [Online]. Available: <http://dx.doi.org/10.1098/rsob.180081>
- [3] D. J. P. Barker et al., "The prenatal origins of lung cancer. ii. the placenta," *Placenta*, 2010, <https://doi.org/10.1016/j.placenta.2010.02.002>.
- [4] D. J. Barker and K. L. Thornburg, "Placental programming of chronic diseases, cancer and lifespan: a review," *Placenta*, vol. 34, no. 10, pp. 841–845, 2013.
- [5] T. Y. Khong, E. E. Mooney, I. Ariel, N. C. M. Balmus, T. K. Boyd, M.-A. Brundler, H. Derricott, M. J. Evans, O. M. Faye-Petersen, J. E. Gillan, A. E. P. Heazell, D. S. Heller, S. M. Jacques, S. Keating, P. Kelehan, A. Maes, E. M. McKay, T. K. Morgan, P. G. J. Nikkels, W. T. Parks, R. W. Redline, I. Scheimberg, M. H. Schoots, N. J. Sebire, A. Timmer, G. Turowski, J. P. van der Voorn, I. van Lijnschoten, and S. J. Gordijn, "Sampling and definitions of placental lesions: Amsterdam placental workshop group consensus statement," *Archives of Pathology & Laboratory Medicine*, vol. 140, no. 7, p. 698–713, Jul. 2016. [Online]. Available: <http://dx.doi.org/10.5858/arpa.2015-0225-CC>
- [6] K. Benirschke, G. J. Burton, and R. N. Baergen, *Pathology of the Human Placenta*, 6th ed. New York: Springer, 2012.
- [7] A. Alansary, K. Kamnitsas, A. Davidson, R. Khlebnikov, M. Rajchl, C. Malamateniou, M. Rutherford, J. V. Hajnal, B. Glocker, D. Rueckert, and B. Kainz, "Fast Fully Automatic Segmentation of the Human Placenta from Motion Corrupted MRI." Springer International Publishing, 2016, p. 589–597. [Online]. Available: [http://dx.doi.org/10.1007/978-3-319-46723-8\\_68](http://dx.doi.org/10.1007/978-3-319-46723-8_68)
- [8] M. Ganesan and V. Shanthi, "Statistical measurement of ultrasound placenta images complicated by gestational diabetes mellitus using segmentation approach," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, pp. 332–343, 10 2011.
- [9] M. Shahedi, J. D. Dormer, A. D. Tensingh Rajan Thanga Kani, Q. N. Do, Y. Xi, M. A. Lewis, A. J. Madhuranthakam, D. M. Twickler, and B. Fei, "Segmentation of uterus and placenta in mr images using a fully convolutional neural network," in *Medical Imaging 2020: Computer Aided Diagnosis*, H. K. Hahn and M. A. Mazurowski, Eds. SPIE, Mar. 2020, p. 59. [Online]. Available: <http://dx.doi.org/10.1117/12.2549873>
- [10] M. Shahedi, C. Y. Spong, J. D. Dormer, Q. N. Do, Y. Xi, M. A. Lewis, C. Herrera, A. J. Madhuranthakam, D. M. Twickler, and B. Fei, "Deep learning-based segmentation of the placenta and uterus on mr images," *Journal of Medical Imaging*, vol. 8, no. 05, Sep. 2021. [Online]. Available: <http://dx.doi.org/10.1117/1.JMI.8.5.054001>
- [11] Y. Chen, C. Wu, Z. Zhang, J. A. Goldstein, A. D. Gernand, and J. Z. Wang, "Automatic morphological characterization of placenta photos with deep learning," in *Proc. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 11764. Cham, Switzerland: Springer, 2019, pp. 487–495, online. Available: [https://doi.org/10.1007/978-3-030-32239-7\\_54](https://doi.org/10.1007/978-3-030-32239-7_54).
- [12] Y. Chen, Z. Zhang, C. Wu, D. Davaasuren, J. A. Goldstein, A. D. Gernand, and J. Z. Wang, "Ai-plax: Ai-based placental assessment and examination using photos," *Computerized Medical Imaging and Graphics*, vol. 84, p. 101744, Sep. 2020. [Online]. Available: <http://dx.doi.org/10.1016/j.compmedimag.2020.101744>
- [13] Z. Zhang, D. Davaasuren, C. Wu, J. A. Goldstein, A. D. Gernand, and J. Z. Wang, "Multi-region saliency-aware learning for cross-domain placenta image segmentation," *Pattern Recognit. Lett.*, vol. 140, pp. 165171, 2020.
- [14] Y. Pan, A. D. Gernand, J. A. Goldstein, L. B. Mithal, D. Mwinyelle, and J. Z. Wang, "Vision-language contrastive learning approach to robust automatic placenta analysis using photographic images," in *Proc. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 13436, 2022, pp. 707–716.
- [15] Y. Pan, T. Cai, M. Mehta, A. D. Gernand, J. A. Goldstein, L. B. Mithal, D. Mwinyelle, K. Gallagher, and J. Z. Wang, "Enhancing automatic placenta analysis through distributional feature recomposition in vision language contrastive learning," in *Proc. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 14393, 2023, pp. 116–126.
- [16] B. A. Kitchenham, *Procedures for Performing Systematic Reviews*, Keele University, 2004. [Online]. Available: <https://www.inf.ufsc.br/aldo.vw/kitchenham.pdf>
- [17] B. S. Lopes, B. Q. T. da Silva, and A. von Wangenheim, "Systematic literature review on computer vision and deep learning for automated screening of macroscopic placental alterations," Technical Report, Universidade Federal de Santa Catarina (UFSC), 2025.
- [18] B. S. Lopes, B. Q. T. da Silva, and A. von Wangenheim, "Placenta Image Capture Protocol in the Delivery Room," Technical Report INCoD/LAPIX.02.2025.B, Universidade Federal de Santa Catarina (UFSC), 2025. [Online]. Available: <https://www.researchgate.net/publication/388499446>
- [19] A. Buslaev, A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Information*, vol. 11, no. 2, p. 125, 2020.
- [20] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. S'anchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361841517301135>
- [21] F. Isensee, P. Jaeger, S. Kohl, J. Petersen, and K. Maier-Hein, "nnunet: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, pp. 1–9, 02 2021.
- [22] A. Estevea, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, Y. Liu, E. Topol, J. Dean, and R. Socher, "Deep learning-enabled medical computer vision," *npj Digital Medicine*, vol. 4, no. 1, Jan. 2021.
- [23] S. Yin, Y. Wang, S. Wang, J. Chang, J. Zhu, and W. Lu, "Tumor segmentation in brain mr images using res-unet with multimodal fusion," in *2025 Asia-Europe Conference on Cybersecurity, Internet of Things and Soft Computing (CITSC)*, 2025, pp. 28–32.
- [24] R. Khanam and M. Hussain, "Yolov11: An overview of the key architectural enhancements," 2024, arXiv preprint arXiv:2410.17725. [Online]. Available: <https://arxiv.org/abs/2410.17725>
- [25] G. Jocher, J. Qiu, and A. Chaurasia, "Ultralytics YOLO 11," <https://github.com/ultralytics/ultralytics>, 2024.
- [26] F. Xiao, "Segmentation of medical biological cell image based on improved unet," in *2024 IEEE 4th Int. Conf. on Electronic Technology, Communication and Information (ICETCI)*. Changchun, China: IEEE, 2024, pp. 67–71.
- [27] Z. Yu, "Pneumonia detection with u-efficientnet," in *2021 IEEE Sixth Int. Conf. on Data Science in Cyberspace (DSC)*. Shenzhen, China: IEEE, 2021, pp. 591–594.
- [28] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A survey on performance metrics for object-detection algorithms," in *2020 International Confer-*

- ence on Systems, Signals and Image Processing (IWSSIP), 2020, pp. 237–242.
- [29] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, “Pytorch: An imperative style, high-performance deep learning library,” in *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
- [30] J. Garland, M. Hu, K. Kesha, C. Glenn, P. Morrow, S. Stables, B. Ondruschka, and R. Tse, “Identifying gross post-mortem organ images using a pre-trained convolutional neural network,” *J. Forensic Sci.*, vol. 66, no. 2, pp. 630–635, 2021
- [31] B. Q. T. da Silva, B. S. Lopes, L. O. Santos, R. P. S. Ribeiro, and A. von Wangenheim, “Towards automated placental screening: Instance segmentation in clinical images,” to appear in *Proc. 38th IEEE Int. Symp. on Computer-Based Medical Systems (CBMS)*, 2025. [Online]. Available: <https://2025.cbms-conference.org/>
- [32] Roboflow, “Roboflow: Annotate, preprocess, and deploy computer vision models,” [Online]. Available: <https://roboflow.com/>
- [33] Roboflow, “Understanding Auto Orientation and EXIF Data,” *Roboflow Blog*, 2023. [Online]. Available: <https://blog.roboflow.com/exif-auto-orientation/>
- [34] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2019. [Online]. Available: <https://arxiv.org/abs/1905.11946>
- [35] Laboratório de Processamento de Imagens e Computação Gráfica (LAPIX), “lapixdl,” *GitHub Repository*. [Online]. Available: <https://github.com/lapix-pucrs/lapixdl>
- [36] Weights & Biases, “Experiment Tracking with W&B,” [Online]. Available: <https://wandb.ai>
- [37] D. S. Nogueira, “EfficientUNet-PyTorch: UNet with EfficientNet backbone,” *GitHub*, [Online]. Available: <https://github.com/danielserrano/efficientunet-pytorch>
- [38] T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Murrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. D. Binder, M. A. Stegmaier, C. Wolf, and U. Koethe, “U-Net: Deep learning for cell counting, detection, and morphometry,” *Nature Methods*, vol. 16, no. 1, pp. 67–70, Jan. 2019.
- [39] J. Maier-Hein, B. M. van Ginneken, S. M. Heimann, M. A. Styner, A. Arbeláez, M. A. Viergever, and A. F. Frangi, “Why rankings of biomedical image analysis competitions should be interpreted with care,” *Nature Communications*, vol. 9, no. 1, pp. 5217, Dec. 2018.
- [40] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [41] K. T. Chang, “Pathological examination of the placenta: raison d’être, clinical relevance and medicolegal utility,” *Singapore Medical Journal*, vol. 50, pp. 1123–1133, 2009.
- [42] N. Mittal, A. Choudhary, A. Singh, V. Kaushik, R. K. Chawla, B. Bhardwaj, and R. Thakur, “A practical guide to placental examination for forensic pathologists,” *Forensic Science, Medicine and Pathology*, vol. 16, pp. 295–312, 2020. [Online]. Available: <https://doi.org/10.1007/s12024-019-00214-2>
- [43] Governo do Distrito Federal, Secretaria de Estado de Saúde do Distrito Federal, Diretoria de Serviços de Urgências, Apoio Diagnóstico e Cirurgias, Gerência de Serviços de Apoio Diagnóstico, “Nota Técnica n. 2/2021 - SES/SAIS/CATES/DUAEC/GEDIAG,” Brasília-DF, 25 Feb. 2021.
- [44] M. C. Del Duque *et al.*, “O erro médico nos tribunais: uma análise das decisões do Tribunal de Justiça da capital brasileira,” *Saúde e Sociedade*, vol. 31, no. 3, 2022. [Online]. Available: <https://doi.org/10.1590/S0104-12902022220144pt>
- [45] T. M. L. C. Rodrigues and A. A. Nunes, “Indenizações em obstetrícia: estudo das decisões do Superior Tribunal de Justiça do Brasil de 2004 a 2014,” *Revista de Direito Sanitário*, vol. 19, no. 1, pp. 121–143, 2018. [Online]. Available: <https://doi.org/10.11606/issn.2316-9044.v19i1p121-143>
- [46] J. Bridge, “Photography in gross specimen pathology,” *Archives of Pathology & Laboratory Medicine*, vol. 133, no. 5, pp. 811–814, 2009.

## **ANEXO A – TOWARDS AUTOMATED PLACENTAL SCREENING: INSTANCE SEGMENTATION IN CLINICAL IMAGES**

**Autores:** Bibiana Quatrin Tiellet da Silva, Beatriz Silva Lopes, Luis Otávio Santos, Rodrigo de Paula e Silva Ribeiro e Aldo von Wangenheim.

**Publicado:** 2025 IEEE 38th International Symposium on Computer-Based Medical Systems (CBMS).

**DOI:** 10.1109/CBMS65348.2025.00016

# Towards Automated Placental Screening: Instance Segmentation in Clinical Images

1<sup>st</sup> Bibiana Quatrin Tiellet da Silva  
*dept. Informatics and Statistics (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brasil  
0000-0001-8774-5643

2<sup>nd</sup> Beatriz Silva Lopes  
*dept. Informatics and Statistics (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brasil  
beatriz.s.lopes@posgrad.ufsc.br

3<sup>rd</sup> Luís Otávio Santos  
*dept. Informatics and Statistics (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brasil  
0000-0002-6660-462X

4<sup>th</sup> Rodrigo de Paula e Silva Ribeiro  
*dept. Informatics and Statistics (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brasil  
0000-0002-6112-8660

5<sup>th</sup> Aldo von Wangenheim  
*dept. Informatics and Statistics (INE)*  
*Federal University of Santa Catarina*  
Florianópolis, Brasil  
0000-0003-4532-1417

**Abstract**—While routine macroscopic placental assessment is performed in all deliveries, anatomopathological examination is reserved for selected cases due to logistical and resource constraints. Although artificial intelligence (AI) has shown promise in medical imaging, its application to postpartum placental screening remains underexplored in clinical settings. This study addresses this gap by developing a deep learning pipeline to improve segmentation of placental structures using macroscopic photographs captured in delivery rooms, with the aim of supporting anatomopathological referral decisions. A proprietary dataset was constructed under both standardized and routine clinical conditions, with annotations curated by pathologists across eight morphological classes. Three architectures—YOLOv11, U-Net with ResNet34, and U-Net with EfficientNet-B0—were evaluated for segmentation and classification performance. Results demonstrated moderate efficacy: YOLOv11 achieved an AP50 of 72% for instance segmentation, while U-Net variants showed higher classification accuracy (up to 85.18%) but lower segmentation precision (Dice scores: 24.70–61.97%). These findings support the feasibility of incorporating AI-based photographic analysis into routine clinical workflows for placental evaluation. Despite preliminary segmentation limitations, this work establishes a foundation for AI-driven placental triage, offering potential to optimize resource allocation by guiding anatomopathological referrals.

**Index Terms**—Placenta, Instance Segmentation, Postpartum Period, Computer Vision, Surgical Pathology.

## I. INTRODUCTION

The placenta plays a crucial role in maternal-fetal interaction, directly influencing pregnancy outcomes and long-term neonatal health [1], [2], [3], [4]. Placental anatomopathological examination is a valuable tool for identifying obstetric complications [5], [6]. However, as most pregnancies proceed without complications, systematic examination of all placentas is logistically and economically unfeasible. Current clinical practice therefore limits this analysis to selected cases, based

on gestational history and macroscopic features observed at delivery [1].

Although Artificial Intelligence (AI) has demonstrated considerable potential for automating medical image analysis, its reported applications in placental pathology appear to be primarily limited to prenatal imaging examinations, such as ultrasound and magnetic resonance [7], [8], [9], [10], and, in the postpartum period, to laboratory-based analyses [11], [12], [13], [14], [15]. This observation suggests a gap in the development of tools for hospital-based placental screening, where preliminary evaluations could assist in decisions regarding anatomopathological examination referrals.

In order to support this investigation, we conducted a Systematic Literature Review (SLR) following Kitchenham’s methodology [16], [17] which confirmed that AI use in this context remains narrowly focused and disconnected from real clinical workflows. To address this gap, this study proposes an innovative solution by developing a deep learning pipeline focused on segmenting and characterizing the macroscopic features of the placenta, aiming to improve the efficiency of examination referrals. For this purpose, we employed a structured image acquisition protocol and a proprietary dataset built from images captured in real clinical environments, specifically in delivery rooms [18].

In this paper, we present the preliminary results of our research, establishing this initial phase as a foundation for future improvements. The next steps involve experimenting with different neural network architectures, fine-tuning hyperparameters, and exploring hybrid approaches, aiming to optimize placental segmentation and integrate the tool into clinical obstetric workflows.

### A. Objectives

The primary objective of this study is to answer the following question:

- Can deep learning models assist in the morphological screening of placentas based on standardized macroscopic photographs captured in real clinical environments?

Addressing this question contributes to a scarcely explored area in computational pathology: the application of AI in macroscopic analysis under clinical conditions. The outcomes of this research may support a more efficient screening process for placental specimens and guide anatomopathological decision-making at the point of care.

## II. MATERIALS AND METHODS

This section outlines the datasets used and describes the methodological framework, including image acquisition procedures, annotation process, data augmentation techniques, and model training strategies.

### A. Ethical Considerations

This research complies with ethical standards for studies involving human subjects, and was approved by the Brazilian Research Ethics Committee (CEP/CONEP) under protocol number CAAE: 86362524.7.0000.0121.

### B. The Placenta

The placenta presents two distinct surfaces: the fetal side, with translucent membranes, vascular arborization, and umbilical cord insertion; and the maternal side, composed of wine-colored spongy tissue with cotyledons separated by thin septa. These gross anatomical features provide diagnostic clues for identifying pathological conditions. Examples of both surfaces and their annotations are shown in Figure 1.

1) *Dataset Description:* This study used a dataset of 156 manually annotated placental images, later expanded to 374 through data augmentation. All images were obtained in delivery rooms of Obstetric Centers and Maternity Hospitals, immediately after delivery. While some were captured under routine hospital conditions, a subset followed a standardized acquisition protocol to ensure consistency in background, lighting, and framing [18]. Data augmentation was performed using the Alumentations library [19]. Horizontal and vertical flips simulated positional variation, optical distortions adjusted depth and perspective, and random contrast, color jitter, and motion blur reflected lighting inconsistencies common in delivery room environments. Affine transformations (rotation, translation, shearing, and zoom) further improved generalization by increasing the model’s ability to adapt to scale and framing variation.

2) *Image Acquisition Protocol:* To standardize the image acquisition, the placenta was positioned on a blue background as described by the protocol of placenta acquisition in the delivery room [18]. This standardization was designed to improve chromatic contrast and align with existing practices in gross pathology and surgical fields. The plate dimensions ensured full accommodation of the placental disc without requiring external scale references, enabling computational inference of relative measurements. All images were obtained with the use of mobile devices (Android A05 and iOS A1778).

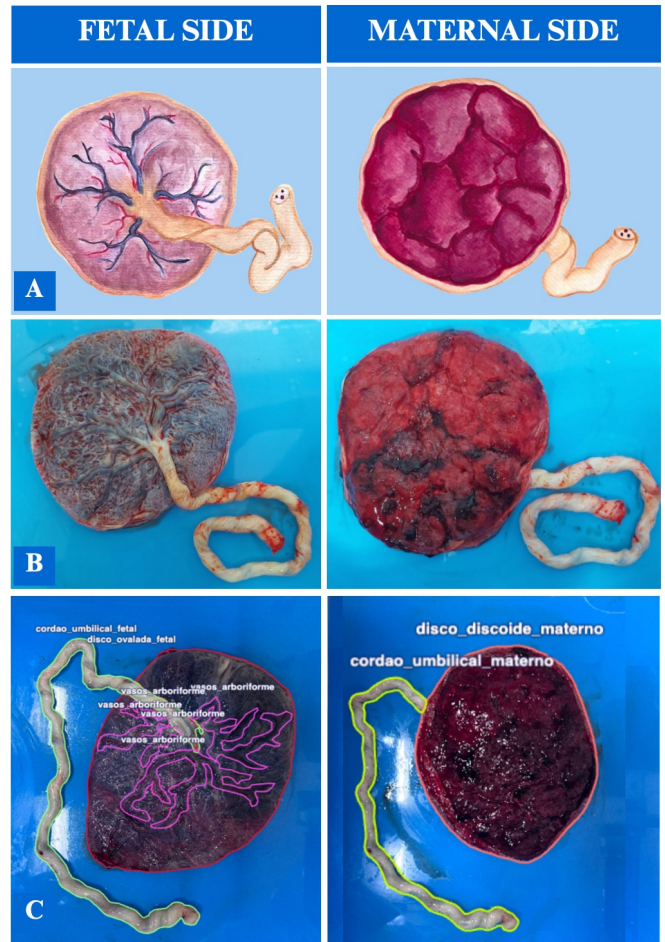


Fig. 1: Anatomical features of the placenta: illustration, photographic capture, and annotated image. Row A: schematic illustration showing the fetal surface (left), with translucent membranes, arboriform vascular pattern, and centrally inserted umbilical cord; and the maternal surface (right), with visible cotyledons separated by thin septa, corresponding to the uterine attachment site. Row B: photographic example of a placenta captured under the standardized imaging protocol, highlighting both surfaces. Row C: annotated version of both side of the placenta image.

3) *Annotation procedures:* To enable supervised training of the AI models, each image was manually annotated with the assistance of a specialized pathologist, ensuring precise curation of the anatomical regions of interest.

Eight classes were defined for visual representation of the morphological characterization of placental structures:

- Fetal umbilical cord – Delimitation of the umbilical cord visible on the fetal surface.
- Maternal umbilical cord – Delimitation of umbilical cord segments visible on the maternal surface.
- Fetal discoid disc – Delimitation of the normal shape of the placental disc on the fetal surface.
- Maternal discoid disc – Delimitation of the normal shape

- of the placental disc on the maternal surface.
- Fetal irregular disc – Delimitation of abnormally shaped placental discs on the fetal surface.
- Maternal irregular disc – Delimitation of abnormally shaped placental discs on the maternal surface.
- Fetal oval-shaped disc – Delimitation of oval-shaped placental discs on the fetal surface.
- Maternal oval-shaped disc – Delimitation of oval-shaped placental discs on the maternal surface.
- Arboriform vessels – Delimitation of vascular structures on the fetal surface.

The final dataset contains the segmentation mask, segmentation points, and label of the object of interest, making it suitable for both semantic and instance segmentation.

### C. Models and Training Procedures

In this study, we evaluated three convolutional neural network (CNN) architectures for the purpose of placental image analysis: YOLO11, U-Net with a ResNet backbone, and U-Net with an efficientnet backbone. We chose these models as they showed state-of-the-art performance in image segmentation tasks on various image modalities, including medical imaging [20]–[23], aiming to assess their effectiveness in the morphological characterization of placental structures. Due to the low amount of data available, the data was split into a training set (80%), validation set (10%), and test set (10%). The total amount of images can be seen in Table I.

1) *YOLO11*: YOLO11 was employed for instance segmentation and object detection. The single-stage architecture predicts bounding boxes and class probabilities in a single forward pass, enabling efficient detection of multiple anatomical structures in macroscopic placental images [24], [25].

2) *U-Net (ResNet34)*: ResNet34 was employed as the encoder backbone within a U-Net architecture for semantic segmentation. This configuration allowed for the pixel-wise delineation of placental regions of interest, such as disc shapes and vascular structures, leveraging ResNet’s feature extraction capabilities with U-Net’s spatial precision [22], [26].

3) *U-Net (EfficientNet-B0)*: This model was explored for its balance between accuracy and computational efficiency. The EfficientNet backbone enhances feature representation while maintaining a lightweight structure, making it a promising approach for semantic segmentation in clinical imaging workflows [22], [27].

4) *Evaluation Metrics*: The performance of the architectures was assessed using metrics widely adopted in the image segmentation field [28]. Intersection over Union (IoU) and the Dice coefficient were applied to quantify the overlap between model predictions and manual annotations for the semantic segmentation models. For instance segmentation, the metrics AP50 and AP50–95 were used to measure average classification and mask precision across different IoU thresholds.

### D. Training Configurations

All models were implemented in Python using high-level libraries, specifically FastAI and PyTorch [29]. All training

was conducted on the university’s NVIDIA DGX H100 server, leveraging its state-of-the-art computational infrastructure to enable efficient and robust data processing. A summary of the training configurations, hyperparameters, and computational details for each evaluated model is presented in Table I.

|                                   | YOLO11        | RESNET34      | EFFICIENT       |
|-----------------------------------|---------------|---------------|-----------------|
| Architecture / Pretrained Model   | Yolov11       | U-Net         | U-Net           |
| Backbone                          | YOLO11n-seg   | ResNet34      | EfficientNet B0 |
| Network Input                     | 640x640       | 512x512       | 512x512         |
| Batch Size                        | 32            | 8             | 8               |
| Epochs                            | 100           | 60            | 60              |
| Optimizer                         | AdamW         | AdamW         | AdamW           |
| Number of images (Train/val/test) | 300 / 37 / 37 | 300 / 37 / 37 | 300 / 37 / 37   |

TABLE I: Training configuration for each model.

## III. RESULTS

The YOLO11 model achieved an average AP50 of 72% and AP50–95 of 67.1%. The ResNet34 U-Net–based approach reached a classification accuracy of 72.86%, with a mask segmentation measured by the mean Dice coefficient of 61.97% and an average IoU of 49.91%. The EfficientNet U-Net model obtained the highest classification accuracy, reaching 85.18%, although its mask segmentation showed an average Dice score of 31.35% and mean IoU of 24.70%.

## IV. DISCUSSION

According to our systematic review [17], previous studies have focused primarily on medical imaging or laboratory environments [7], [8], [9], [11], [12]. In contrast, this study used photographs captured directly in the delivery room under semi-controlled conditions, bringing the computational approach closer to real-world clinical practice. A subset of these images followed a standardized acquisition protocol, while others reflected the variability commonly found in hospital settings.

Despite satisfactory performance in class classification, the segmentation metrics were suboptimal, suggesting the need for more annotated images and further refinement of the dataset and model hyperparameters. A closer examination of the dataset reveals that the challenges in segmentation are particularly associated with anatomical regions exhibiting complex morphology. Figure 2 shows examples of visual artifacts, such as blood overlap and inadequate lighting. Figure 3 illustrates how variations in capture conditions, such as

lighting, background, and framing, can affect both annotation quality and model performance.

The fetal surface contains a tortuous vascular arrangement characterized by high inter-individual variability and poorly defined boundaries, complicating the accurate mapping of vessel trajectories in two-dimensional images. Additionally, overlapping tissues, uneven lighting, and surface blood residues further impair the annotator’s ability to consistently delineate structures. The photographic protocol [18] serves as a proposed strategy to mitigate these issues through standardization of lighting, framing, and background during image acquisition, findings frequently highlighted in the literature: image acquisition standardization plays a critical role in the performance of computer vision models [13], [14], [30].

YOLO11 showed average performance in both classification and segmentation (AP metric). U-Net-based models ResNet34 and EfficientNet-B0 demonstrated good overall classification but poor segmentation precision. These limitations are consistent with observations reported in related studies of semantic segmentation [11], [12].

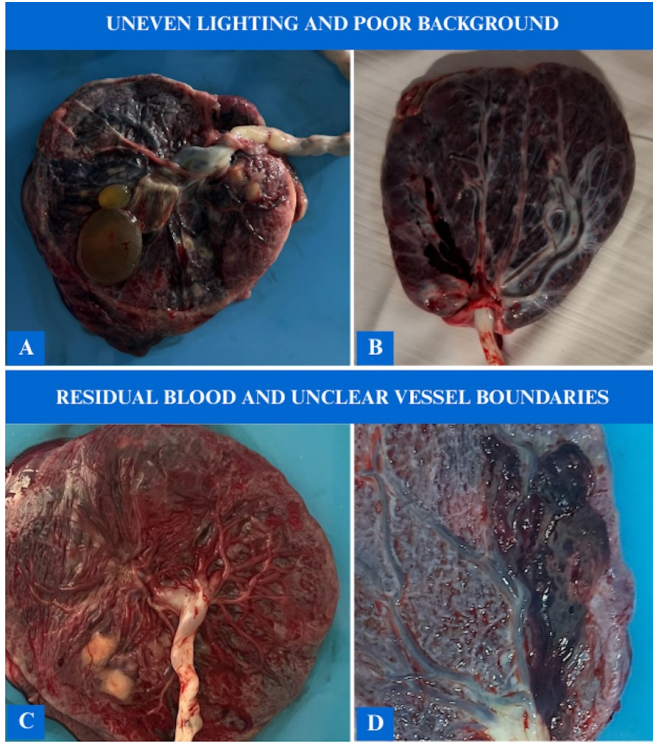


Fig. 2: Examples of placental images presenting visual challenges for segmentation. (A) and (B) display uneven lighting and non-standard backgrounds. (C) and (D) show residual blood and unclear vessel boundaries. These visual artifacts can compromise both manual annotation and model performance.

## V. LIMITATIONS

While this study offers preliminary insights into AI-driven placental screening, some limitations must be acknowledged.

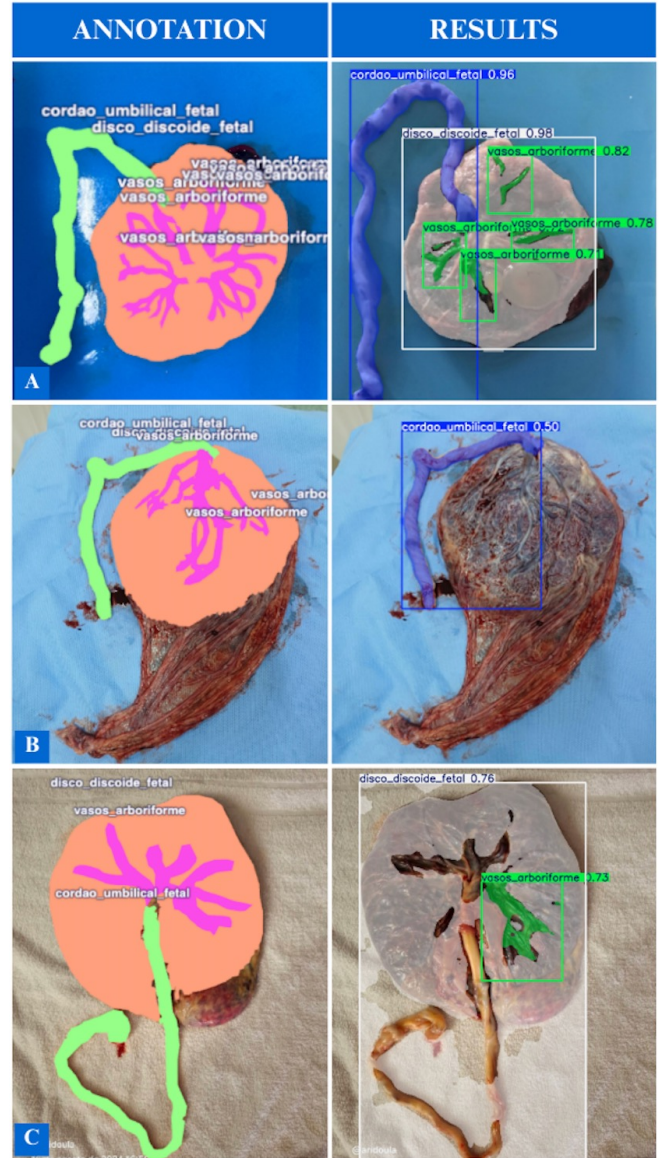


Fig. 3: Annotation and model prediction across different image capture conditions. Row A: (left) Annotated image of Placenta captured with the standardized protocol; (right) segmentation results generated by the YOLO11 model. Row B: (left) Annotated image captured in a hospital setting without standardization, using the typical blue background of surgical units; (right) corresponding YOLO11 prediction. Row C: (left) Annotated image provided by a doula, captured under non-standard conditions with a white background; (right) YOLO11 prediction.

- The limited size (374 images) and class imbalance of the dataset restrict model generalizability, particularly for rare morphological anomalies. The next step should focus on more diverse datasets to improve segmentation fidelity.
- Although a standardized acquisition protocol was partially implemented, inconsistent adherence in real-world

settings underscores the need for stricter image acquisition guidelines to minimize visual artifacts (e.g., blood residues, uneven backgrounds). Future iterations should compare a completed standardized dataset.

- The evaluation was conducted exclusively in academic environments, leaving the models' real-world clinical utility untested. Integration into hospital workflows and direct benchmarking against existing tools are essential next steps to assess practical impact.
- Architectural exploration focused on three models without addressing new transformer-based models, which could potentially improve fine segmentation via attention mechanisms.

Despite these challenges, this work establishes a foundational framework for AI-assisted placental triage. Addressing these limitations through expanded datasets, evaluation of newer architectures (e.g., transformers), and clinical validation will be essential to bridge the gap between computational research and point-of-care deployment.

## VI. CONCLUSION

Our study represents an initial round of experimentation exploring the use of AI for placental screening in obstetric settings.

Preliminary results suggest that deep learning models can assist in the morphological analysis of the placenta and inform decisions regarding anatomopathological examination requests with great potential to increase performance and precision when supported by standardized imaging protocols and well-annotated datasets.

In current practice, anatomopathological examination of the placenta is performed based on clinical conditions or macroscopic findings suggestive of disease. This decision relies on visual inspection by the obstetrician or attending physician in the delivery room. However, due to the subjective nature of this assessment and the impracticality of performing the examination on all placentas, the process may result in unnecessary referrals or the underuse of pathological analysis. Thus, although still preliminary, the proposed approach demonstrates potential to optimize resource allocation in placental triage, assisting physicians in the delivery room.

As healthcare systems increasingly adopt AI-driven pipelines, model selection must prioritize architectures that balance precision and generalizability. Our work bridges AI innovation with clinical pragmatism, offering a framework to optimize placental screening workflow. Future directions include expanding the dataset to improve model generalization in domain-specific applications, further bridging the gap between computational research and clinical deployment. Subsequent steps involve validation in controlled laboratory environments and field testing in real-world clinical settings, with the aim of integrating the system into routine workflows, potentially through dedicated mobile applications.

## ACKNOWLEDGMENT

We thank all the postpartum women who consented to the photographic documentation of their placentas.

This work was performed with computational resources provided by Fundação de Amparo a Pesquisa e Inovação do Estado de Santa Catarina (FAPESC), grant nr. 2024TR000090.

## REFERENCES

- [1] D. J. Roberts, "Placental pathology, a survival guide," *Archives of Pathology & Laboratory Medicine*, vol. 132, no. 4, p. 641–651, Apr. 2008. [Online]. Available: <http://dx.doi.org/10.5858/2008-132-641-PPASG>
- [2] V. Costanzo, A. Bardelli, S. Siena, and S. Abrignani, "Exploring the links between cancer and placenta development," *Open Biology*, vol. 8, no. 6, Jun. 2018. [Online]. Available: <http://dx.doi.org/10.1098/rsob.180081>
- [3] D. J. P. Barker *et al.*, "The prenatal origins of lung cancer. ii. the placenta," *Placenta*, 2010, <https://doi.org/10.1016/j.placenta.2010.02.002>.
- [4] D. J. Barker and K. L. Thornburg, "Placental programming of chronic diseases, cancer and lifespan: a review," *Placenta*, vol. 34, no. 10, pp. 841–845, 2013.
- [5] T. Y. Khong, E. E. Mooney, I. Ariel, N. C. M. Balmus, T. K. Boyd, M.-A. Brundler, H. Derricott, M. J. Evans, O. M. Faye-Petersen, J. E. Gillan, A. E. P. Heazell, D. S. Heller, S. M. Jacques, S. Keating, P. Kelehan, A. Maes, E. M. McKay, T. K. Morgan, P. G. J. Nikkels, W. T. Parks, R. W. Redline, I. Scheimberg, M. H. Schoots, N. J. Sebire, A. Timmer, G. Turowski, J. P. van der Voorn, I. van Lijnschoten, and S. J. Gordijn, "Sampling and definitions of placental lesions: Amsterdam placental workshop group consensus statement," *Archives of Pathology & Laboratory Medicine*, vol. 140, no. 7, p. 698–713, Jul. 2016. [Online]. Available: <http://dx.doi.org/10.5858/arpa.2015-0225-CC>
- [6] K. Benirschke, G. J. Burton, and R. N. Baergen, *Pathology of the Human Placenta*, 6th ed. New York: Springer, 2012.
- [7] A. Alansary, K. Kamnitsas, A. Davidson, R. Khlebnikov, M. Rajchl, C. Malamateniou, M. Rutherford, J. V. Hajnal, B. Glocker, D. Rueckert, and B. Kainz, *Fast Fully Automatic Segmentation of the Human Placenta from Motion Corrupted MRI*. Springer International Publishing, 2016, p. 589–597. [Online]. Available: [http://dx.doi.org/10.1007/978-3-319-46723-8\\_68](http://dx.doi.org/10.1007/978-3-319-46723-8_68)
- [8] M. Ganesan and V. Shanthi, "Statistical measurement of ultrasound placenta images complicated by gestational diabetes mellitus using segmentation approach," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, pp. 332–343, 10 2011.
- [9] M. Shahedi, J. D. Dormer, A. D. Tensingh Rajan Thanga Kani, Q. N. Do, Y. Xi, M. A. Lewis, A. J. Madhuranthakam, D. M. Twickler, and B. Fei, "Segmentation of uterus and placenta in mr images using a fully convolutional neural network," in *Medical Imaging 2020: Computer-Aided Diagnosis*, H. K. Hahn and M. A. Mazurowski, Eds. SPIE, Mar. 2020, p. 59. [Online]. Available: <http://dx.doi.org/10.1117/12.2549873>
- [10] M. Shahedi, C. Y. Spong, J. D. Dormer, Q. N. Do, Y. Xi, M. A. Lewis, C. Herrera, A. J. Madhuranthakam, D. M. Twickler, and B. Fei, "Deep learning-based segmentation of the placenta and uterus on mr images," *Journal of Medical Imaging*, vol. 8, no. 05, Sep. 2021. [Online]. Available: <http://dx.doi.org/10.1117/1.JMI.8.5.054001>
- [11] Y. Chen, C. Wu, Z. Zhang, J. A. Goldstein, A. D. Gernand, and J. Z. Wang, "Automatic morphological characterization of placenta photos with deep learning," in *Proc. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 11764. Cham, Switzerland: Springer, 2019, pp. 487–495, online. Available: [https://doi.org/10.1007/978-3-030-32239-7\\_54](https://doi.org/10.1007/978-3-030-32239-7_54).
- [12] Y. Chen, Z. Zhang, C. Wu, D. Davaasuren, J. A. Goldstein, A. D. Gernand, and J. Z. Wang, "Ai-plax: Ai-based placental assessment and examination using photos," *Computerized Medical Imaging and Graphics*, vol. 84, p. 101744, Sep. 2020. [Online]. Available: <http://dx.doi.org/10.1016/j.compmedimag.2020.101744>
- [13] Z. Zhang, D. Davaasuren, C. Wu, J. A. Goldstein, A. D. Gernand, and J. Z. Wang, "Multi-region saliency-aware learning for cross-domain placenta image segmentation," *Pattern Recognit. Lett.*, vol. 140, pp. 165–171, 2020.

- [14] Y. Pan, A. D. Gernand, J. A. Goldstein, L. B. Mithal, D. Mwinyelle, and J. Z. Wang, "Vision-language contrastive learning approach to robust automatic placenta analysis using photographic images," in *Proc. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 13436, 2022, pp. 707–716.
- [15] Y. Pan, T. Cai, M. Mehta, A. D. Gernand, J. A. Goldstein, L. B. Mithal, D. Mwinyelle, K. Gallagher, and J. Z. Wang, "Enhancing automatic placenta analysis through distributional feature recomposition in vision-language contrastive learning," in *Proc. Med. Image Comput. Comput. Assist. Interv. (MICCAI)*, vol. 14393, 2023, pp. 116–126.
- [16] B. A. Kitchenham, *Procedures for Performing Systematic Reviews*. Keele University, 2004. [Online]. Available: <https://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf>
- [17] B. S. Lopes, B. Q. T. da Silva, and A. V. Wangenheim, "Systematic literature review on computer vision and deep learning for automated screening of macroscopic placental alterations," *Universidade Federal de Santa Catarina*, 2025.
- [18] B. S. Lopes, B. Q. T. da Silva, and A. von Wangenheim, "Protocolo de captura de imagens de placentas na sala de parto," INCoD/UFSC, Florianópolis, Relatório Técnico INCoD/LAPIX.02.2025.B, 2025. [Online]. Available: [https://www.researchgate.net/publication/388499446\\_PLACENTA\\_IMAGE\\_CAPTURE\\_PROTOCOL\\_IN\\_THE\\_DELIVERY\\_ROOMPROTOCOLO\\_DE\\_CAPTURA\\_DE\\_IMAGENS\\_DE\\_PLACENTAS\\_NA\\_SALA\\_DE\\_PARTO](https://www.researchgate.net/publication/388499446_PLACENTA_IMAGE_CAPTURE_PROTOCOL_IN_THE_DELIVERY_ROOMPROTOCOLO_DE_CAPTURA_DE_IMAGENS_DE_PLACENTAS_NA_SALA_DE_PARTO)
- [19] A. Buslaev, A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," *Information*, vol. 11, no. 2, p. 125, 2020.
- [20] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1361841517301135>
- [21] F. Isensee, P. Jaeger, S. Kohl, J. Petersen, and K. Maier-Hein, "nnunet: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, pp. 1–9, 02 2021.
- [22] A. Esteva, K. Chou, S. Yeung, N. Naik, A. Madani, A. Mottaghi, Y. Liu, E. Topol, J. Dean, and R. Socher, "Deep learning-enabled medical computer vision," *npj Digital Medicine*, vol. 4, no. 1, Jan. 2021.
- [23] S. Yin, Y. Wang, S. Wang, J. Chang, J. Zhu, and W. Lu, "Tumor segmentation in brain mr images using res-unet with multimodal fusion," in *2025 Asia-Europe Conference on Cybersecurity, Internet of Things and Soft Computing (CITSC)*, 2025, pp. 28–32.
- [24] R. Khanam and M. Hussain, "Yolov11: An overview of the key architectural enhancements," 2024, arXiv preprint arXiv:2410.17725. [Online]. Available: <https://arxiv.org/abs/2410.17725>
- [25] G. Jocher, J. Qiu, and A. Chaurasia, "Ultralytics YOLO 11," <https://github.com/ultralytics/ultralytics>, 2024.
- [26] F. Xiao, "Segmentation of medical biological cell image based on improved unet," in *2024 IEEE 4th Int. Conf. on Electronic Technology, Communication and Information (ICETCI)*. Changchun, China: IEEE, 2024, pp. 67–71.
- [27] Z. Yu, "Pneumonia detection with u-efficientnet," in *2021 IEEE Sixth Int. Conf. on Data Science in Cyberspace (DSC)*. Shenzhen, China: IEEE, 2021, pp. 591–594.
- [28] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A survey on performance metrics for object-detection algorithms," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2020, pp. 237–242.
- [29] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshin, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
- [30] J. Garland, M. Hu, K. Kesha, C. Glenn, P. Morrow, S. Stables, B. Ondruschka, and R. Tse, "Identifying gross post-mortem organ images using a pre-trained convolutional neural network," *J. Forensic Sci.*, vol. 66, no. 2, pp. 630–635, 2021.