

**Universidade do Grande Rio “Prof. José de Souza Herdy”**

**UNIGRANRIO**

**Lilith Antoniely Lima Porto**

**PADRONIZAÇÃO DE UM ENSAIO DE MICROSCOPIA AUTOMATIZADA PARA  
AVALIAR TOXICIDADE DE COMPOSTOS QUÍMICOS SOBRE EMBRIÕES DE  
*ZEBRAFISH***

**Duque de Caxias - RJ**

**2025**

**Universidade do Grande Rio “Prof. José de Souza Herdy”**

**UNIGRANRIO**

**Lilith Antoniely Lima Porto**

Trabalho de Conclusão de Curso apresentado à Universidade do Grande Rio “Prof. José de Souza Herdy”, como requisito parcial para a obtenção do título de Bacharel em Biomedicina.

Orientador: Dr. Rafael Ferreira Dantas e Dra. Giuliana Viegas Schirato.

**Duque de Caxias - RJ**

**2025**

**PADRONIZAÇÃO DE UM ENSAIO DE MICROSCOPIA AUTOMATIZADA PARA  
AVALIAR TOXICIDADE DE COMPOSTOS QUÍMICOS SOBRE EMBRIÕES DE  
*ZEBRAFISH***

Trabalho de Conclusão de Curso apresentado à  
Universidade do Grande Rio “Prof. José de Souza  
Herdy”, como requisito parcial para a obtenção do título  
de Bacharel em Biomedicina.

Orientador: Dr. Rafael Ferreira Dantas e Dra. Giuliana  
Viegas Schirato.

Aprovado em:

Duque de Caxias, \_\_ de \_\_\_\_\_ de 2025

BANCA EXAMINADORA



---

Prof. Dr. Rafael Ferreira Dantas



---

Prof. Dra. Renata Cristine Manfrinato Reis

Documento assinado digitalmente  
 **THAIS MORAIS DE BRITO**  
Data: 26/06/2025 18:11:49-0300  
Verifique em <https://validar.it.gov.br>

---

Prof. Dra. Thais Morais de Brito

**Duque de Caxias - RJ**

**2025**

### **Agradecimentos**

Em primeiro lugar, agradeço a Deus por todas as maravilhas que Ele tem realizado em minha vida, e por cada pessoa que Ele colocou em meu caminho. À minha mãe, Antonia Souza, e o meu noivo, Andre Passos, expresso minha profunda gratidão pelo apoio incondicional, por acreditarem no meu potencial e por investirem com tanto amor no meu sucesso. Ao grupo LABECFAR, agradeço pela oportunidade de fazer parte dessa equipe, por abrir a porta fundamental da minha construção acadêmica. Agradeço também ao Laboratório de Fisiologia do INCQS pela colaboração com este projeto, por cada apoio e ensinamento, em especial aos Doutores Renata Medeiros, Magno Magalhães e Thais Morais. Meus sinceros agradecimentos aos meus orientadores, Dr. Rafael Ferreira Dantas e Dra. Giuliana Viegas Schirato, que são como pais para mim, sempre exercendo um papel essencial como meus mentores acadêmicos. A paciência, dedicação e incentivo de ambos foram indispensáveis para o meu desenvolvimento acadêmico, assim como sua presença constante em momentos especiais da minha jornada. Estendo a minha gratidão aos meus amigos Beatriz Ferreira de Sá e Jordano Ferreira Reis, pelo apoio constante, companheirismo e ajuda em diversos aspectos da minha vida. A todas as pessoas que contribuíram para o meu desenvolvimento ao longo desse caminho, meu sincero muito obrigada. Estarei sempre ao lado de vocês para celebrar suas conquistas, assim como vocês estiveram presentes nas minhas.

## Lista de Figuras

Figura 1 - Fase embrionária (1h – 96hpf), fase larval (120hpf) e adulta (90 dias) do zebrafish. .....	4
Figura 2 – Alterações morfológicas/funcionais usados pelo teste 236 da OECD para avaliar a toxicidade aguda em embriões de zebrafish de 96 hpf. ....	6
Figura 3 - Fluxo de trabalho proposto para o ensaio. ....	11
Figura 4 - Imagem original e máscaras binárias do corpo inteiro, saco vitelino e área pericárdica. ....	13
Figura 5 - Esquema geral de validação cruzada aninhada com 3 loops externos e 2 internos. ....	15
Figura 6 - Principais fenótipos observados em embriões de zebrafish tratados ou não com os compostos teste. ....	17
Figura 7 - Resultados obtidos após o treinamento do modelo de segmentação para o corpo inteiro de embrião de zebrafish. ....	18
Figura 8 - Resultados obtidos após o treinamento do modelo de segmentação para o saco vitelino do embrião de zebrafish. ....	19
Figura 9 - Resultados obtidos após o treinamento do modelo de segmentação para a área pericárdica do embrião de zebrafish. ....	19
Figura 10 - Avaliação dos parâmetros “area” e “form fator” nos diferentes fenótipos de zebrafish. ....	21
Figura 11 - Avaliação da área do saco vitelino (A) e da área pericárdica (B) em embriões de zebrafish com fenótipos "normal" e "alterado". ....	22
Figura 12 – Performance do classificador baseado em IA. ....	23

## SUMÁRIO

1	INTRODUÇÃO .....	4
1.1	<i>Zebrafish</i> e suas aplicações na descoberta de fármacos .....	4
1.2	Avaliação de toxicidade utilizando o <i>zebrafish</i> .....	6
1.1.2	Uso do HCS e sua implementação no modelo <i>zebrafish</i> .....	7
2	METODOLOGIA .....	10
2.1	Compostos químicos .....	10
2.2	<i>Zebrafish</i> .....	10
2.3	Fluxo de trabalho proposto para o ensaio de toxicidade baseado em HCA .....	10
2.4	Ensaio de toxicidade por microscopia de alto conteúdo .....	11
2.4.1	Preparação dos ensaios com embriões de <i>zebrafish</i> .....	11
2.4.2	Aquisição das imagens por HCA .....	12
2.4.3	Análise das imagens .....	12
2.4.3.1	Treinamento de modelos de inteligência artificial para segmentação automática dos embriões de <i>zebrafish</i> e de seus órgãos .....	12
2.4.3.2	Extração automática de parâmetros fenotípicos do <i>zebrafish</i> .....	15
2.4.3.3	Análise dos fenótipos .....	15
2.4.3.4	Classificador baseado em IA .....	16
3	RESULTADOS .....	17
3.1	Análises dos fenótipos do <i>zebrafish</i> .....	17
3.2	Segmentação das imagens por IA .....	18
3.3	Avaliação do efeito de compostos teste sobre embriões de <i>zebrafish</i> .....	20
3.4	Modelo de classificação gerado por IA .....	22
4	DISCUSSÃO .....	23
5	CONCLUSÃO .....	28
	REFERÊNCIAS .....	29
	ANEXOS .....	34

# **PADRONIZAÇÃO DE UM ENSAIO DE MICROSCOPIA AUTOMATIZADA PARA AVALIAR TOXICIDADE DE COMPOSTOS QUÍMICOS SOBRE EMBRIÕES DE *ZEBRAFISH***

Lilith Antoniely Lima Porto<sup>1</sup>

Rafael Ferreira Dantas<sup>2</sup>

Giuliana Viegas Schirato<sup>3</sup>

## **RESUMO**

O *zebrafish* (*Danio rerio*) é um modelo experimental amplamente utilizado em pesquisas biomédicas, como na descoberta de fármacos. Neste trabalho, iniciamos a padronização de um novo ensaio automatizado de análise de alto conteúdo para avaliar a toxicidade de compostos químicos sobre embriões *wild-type* de *zebrafish*. Para isso, inicialmente testamos compostos químicos reconhecidamente tóxicos sobre o *zebrafish* para avaliar os seus efeitos em embriões com 96 horas pós fertilização. Embora os experimentos não tenham sido validados segundo critérios do teste 236 de toxicidade aguda da OECD, as imagens geradas nesses experimentos foram úteis no desenvolvimento de três etapas operacionais do ensaio: segmentação por inteligência artificial (rede neural convolucional, fast.ai) dos embriões e de seus órgãos (saco vitelino e área pericárdica), quantificação de seus parâmetros morfológicos (CellProfiler) e classificação por IA (XGBoost) dos indivíduos em fenótipos gerais (“normal”, “não eclodido” e “alterado”). Após o treinamento, os modelos de IA mostraram-se adequados para a segmentação do corpo do *zebrafish* nas imagens (Dice = 0,95), mas mostraram menor desempenho para o saco vitelino (Dice = 0,8) e área pericárdica (Dice = 0,25), esses últimos apresentando possíveis sinais de *overfitting*. A análise fenotípica monoparamétrica revelou que o parâmetro morfológico “form factor” mostrou-se eficaz na detecção de embriões não eclodidos. Já na análise multiparamétrica, o classificador fenotípico mostrou uma boa performance (MCC = 0,84), apresentando alta capacidade de identificar embriões com fenótipo “normal” e “não eclodido” (>90% acerto), mas menor desempenho para a classe “alterado” (65-78%). Novos experimentos são necessários para finalizar a padronização do ensaio.

**Palavras-chave:** *Zebrafish*; embriões de *zebrafish*; ensaio de toxicidade; análise de alto conteúdo (HCA); inteligência artificial (IA).

## ABSTRACT

The zebrafish (*Danio rerio*) is a broadly used experimental model in biomedical research, as in drug discovery. In this work, we initiated standardization of a new automated high-content analysis assay to evaluate the toxicity of chemical compounds on wild-type zebrafish embryos. To this end, we initially tested chemical compounds known to be toxic to zebrafish to evaluate their effects on embryos at 96 hours after fertilizations. Although the experiments were not validated according to the OECD acute toxicity test 236 criteria, the images generated in these experiments were useful in the developments of three operational steps of the assay: segmentation by artificial intelligence (convolutional neural network, fast.ai) of the embryos and their organs (yolk sac and pericardial area), quantification of their morphological parameters (CellProfiler) and AI classification (XGBoost) of the individuals into general phenotypes (“normal”, “not hatched” and “altered”). After training, the AI models were adequate at segmenting the zebrafish body in the images (Dice = 0.95), but showed lower performance for the yolk sac (Dice = 0.8) and pericardial area (Dice = 0.25), the latter presenting possible signs of overfitting. The monoparametric phenotypic analysis revealed that the morphological parameter “form factor” was effective in detecting unhatched embryos. In the multiparametric analysis, the phenotypic classifier showed good performance (MCC = 0.84), presenting a high capacity to identify embryos with “normal” and “not hatched” phenotypes (>90% accuracy), but lower performance for the “altered” class (65-78%). New experiments are needed to finalize the standardization of the assay.

**Keywords:** Zebrafish; zebrafish embryos; toxicity assay; high content analysis; artificial intelligence (AI).

<sup>1</sup>Graduando(a) do curso de Biomedicina na Universidade do Grande Rio – UNIGRANRIO.

<sup>2</sup>Biomédico pela UNIRIO. Mestre e doutor em Biociência (UERJ)

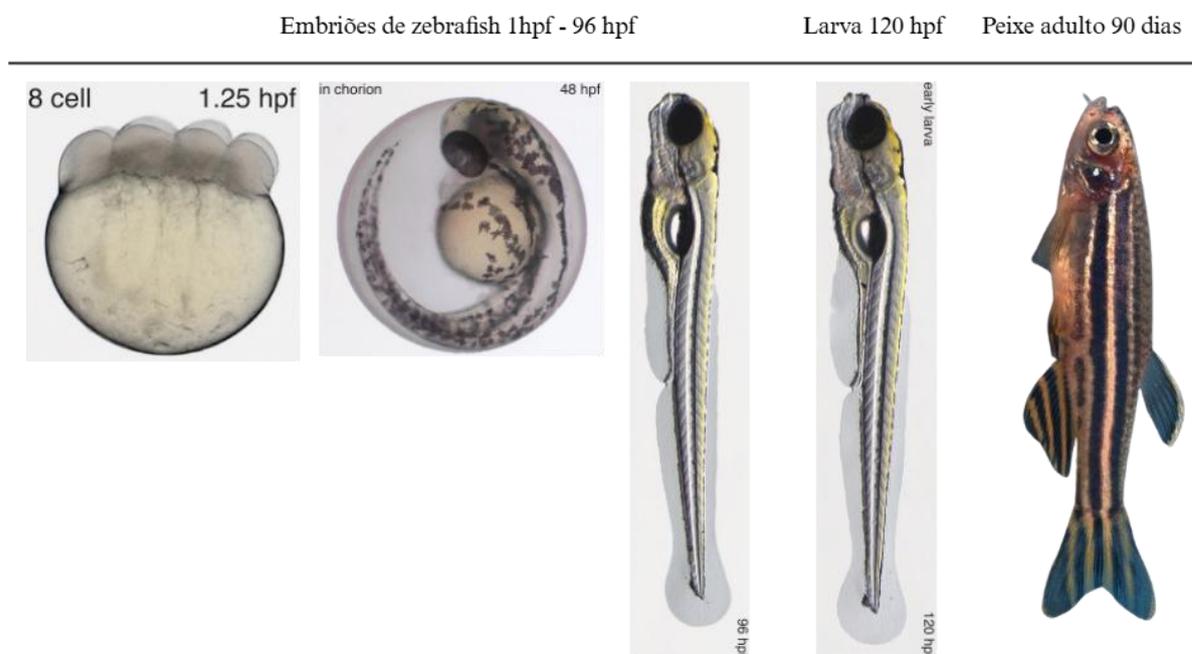
<sup>3</sup>Veterinária pela UFRP. Mestre em ciências Biológicas e doutora em Medicina veterinária (UFRP)

## 1 INTRODUÇÃO

### 1.1 *Zebrafish* e suas aplicações na descoberta de fármacos

O teleosteo da espécie *Danio rerio*, comumente conhecido como *zebrafish* (“peixe-zebra”), é um peixe tropical originário do sudeste da Ásia que começou a ser estudado em laboratório há mais de 40 anos atrás como um organismo modelo no campo da biologia do desenvolvimento (Patton; Zon; Langenau, 2021) . Desde então ele tem contribuído em diversas áreas, como na genética, embriologia, toxicologia ambiental e na descoberta de fármacos (Cassar *et al.*, 2020; Patton; Zon; Langenau, 2021). Parte do seu sucesso é atribuída as seguintes características: pequeno tamanho do corpo ao longo de todas as fases de seu desenvolvimento (embriões com aproximadamente 0,7 mm de diâmetro, larvas com 2-4 mm de comprimento e adultos medindo até 4 cm de comprimento) (Figura 1). Outros fatores que contribuem para o seu uso em laboratório incluem: capacidade de realizar a fertilização *ex vivo*, geração de centenas de embriões a partir de um casal reprodutor e a transparência do embrião, a qual permite a visualização de órgãos internos, sistemas (ex: cardiovascular) e até mesmo células individuais (Miscovic; Rotstein; Wen, 2012; Otterstrom *et al.*, 2022; Patton; Zon; Langenau, 2021).

Figura 1 - Fase embrionária (1h – 96hpf), fase larval (120hpf) e adulta (90 dias) do *zebrafish*. hpf – horas após a fertilização.



Fonte: Adaptado de: Canny *et al.*, .2009

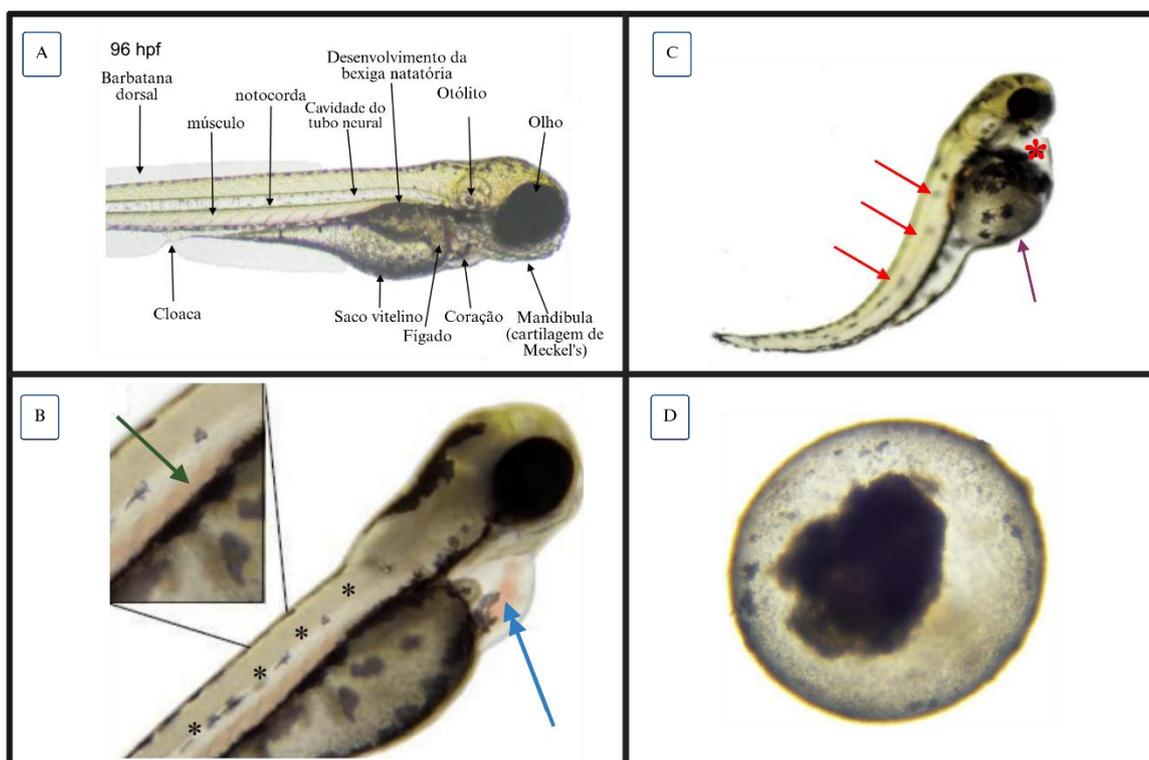
O *zebrafish* compartilha 70% de seu genoma com os humanos e mais de 80% de proteínas relacionadas a doenças são conservadas. Desta forma, compostos ativos em seres humanos podem possuir o mesmo alvo no *zebrafish* (Patton; Zon; Langenau, 2021). Além disso, ele pode ser manipulado geneticamente através de técnicas como CRISPR, permitindo a geração de modelos mais geneticamente precisos (Cully, 2019; Patton; Zon; Langenau, 2021). Isto torna o *zebrafish* um organismo muito útil no estudo do funcionamento e desenvolvimento de sistemas humanos, assim como na elucidação da etiologia, progressão e processos que levam a resolução de doenças. Atualmente existem diversos modelos de *zebrafish* para doenças humanas, como câncer, doenças hepáticas e cardíacas, e distúrbios hematológicos e comportamentais (Patton; Zon; Langenau, 2021). Modelos de *zebrafish* transgênicos também têm sido amplamente usados em campanhas de *screening* fenotípico com o intuito de identificar compostos com potencial farmacológico e avaliar sua toxicidade (Garcia; Noyes; Tanguay, 2016; Miscevic; Rotstein; Wen, 2012). Em relação a outros modelos pré-clínicos, como células humanas e camundongos, o *zebrafish* se destaca por ser um organismo vertebrado (i.e. possui a complexidade de um ser vivo vertebrado completo) e poder ser usado numa vazão muito superior aos modelos murinos, em função do seu menor custo de manutenção e maior taxa de maturação e reprodução. Além disso, em função da transparência e desenvolvimento externo do embrião, é possível usar técnicas não invasivas de imageamento para avaliar o efeito de manipulação genética ou de tratamentos com compostos químicos, o que minimiza o sofrimento e estresse do animal (Cully, 2019; Garcia; Noyes; Tanguay, 2016).

A maioria dos estudos toxicológicos com *zebrafish* tem focado em contaminantes ambientais, mas muito tem se avançado também no campo da toxicologia farmacológica. De acordo com um levantamento realizado entre 2007 e 2018 houve um aumento de cerca de três vezes no número de publicações científicas com *zebrafish* e um incremento de quatro vezes em estudos que usaram este modelo para avaliar toxicidade (Cassar *et al.*, 2020). Dentre os ensaios mais relevantes incluem aqueles que avaliam a cardio, neuro, hepato e nefrotoxicidade, uma vez que tais efeitos deletérios são comumente observados em ensaios clínicos humanos e são responsáveis pela remoção subsequente de muitos fármacos do mercado (Bauer; Mally; Liedtke, 2021).

## 1.2 Avaliação de toxicidade utilizando o *zebrafish*

Tradicionalmente, a avaliação da toxicidade de compostos químicos sobre o *zebrafish* é realizada por microscopia óptica e requer um especialista para representar as alterações observadas num *score* fenotípico (Jeanray *et al.*, 2015). A Organização para a Cooperação e Desenvolvimento Econômico - *Organization for Economic Co-operation and Development* (OECD), por exemplo, recomenda uma série de protocolos para avaliar a toxicidade em *zebrafish* (Padilla; Glaberman, 2020). Dentre eles, o teste de toxicidade aguda em embriões de peixes (teste 236), o qual avalia, por microscopia de campo claro, indicadores de letalidade (ex: ausência de batimentos cardíacos, coagulação de ovos fertilizados, ausência de somitos, não descolamento da cauda) do embrião ao longo de 96h de incubação com o composto teste (OECD, 2013). Durante a análise também são registradas outras alterações morfológicas que indicam algum grau de toxicidade, como o aumento da área pericárdica e do saco vitelino (Figura 2). Como esperado, estes métodos tendem a ser lentos, subjetivos e semiquantitativos (Jeanray *et al.*, 2015; Zhang *et al.*, 2021).

Figura 2 – Alterações morfológicas/funcionais usados pelo teste 236 da OECD para avaliar a toxicidade aguda em embriões de *zebrafish* de 96 hpf. Fenótipo normal de um embrião (A). Embriões com fenótipos alterados (B-D). Nessas imagens são mostradas algumas alterações morfológicas/funcionais que indicam algum grau de toxicidade ou letalidade, como a ausência de somitos (setas vermelhas), imobilidade de células sanguíneas (seta verde), perda de batimento cardíaco (seta azul), edema da área pericárdica (\* vermelho) e saco vitelino (seta roxa, imagem C), e coagulação do embrião (D).



Fonte: Adaptado de: Hellfeld *et al.*, 2020; OECD, 2013

Na tentativa de contornar as limitações da microscopia tradicional, algumas soluções mais automatizadas têm sido propostas. Na maioria dos ensaios a análise é feita com as formas embrionárias e larvais do *zebrafish* distribuídas em microplacas de 96-384 poços ou sistemas de microfluidos (Chapela *et al.*, 2019; Chen *et al.*, 2022; Khalili; Rezai, 2019; Otterstrom *et al.*, 2022; Renquist *et al.*, 2013; Yozzo *et al.*, 2013; Zhu *et al.*, 2020). Dentre as soluções desenvolvidas até o momento, incluem alguns métodos baseados em leitores de microplacas que quantificam, por absorvância, fluorescência ou outros tipos de luminescência, o efeito de compostos sobre o metabolismo energético, produção de espécies reativas de oxigênio, ativação de caspases e tempo de trânsito intestinal (Cassar; Huang; Cole, 2015; Duan *et al.*, 2016; Reid; D'Aquila; Biga, 2018; Seng; Zhang; McGrath, 2016). No entanto, as soluções mais comuns empregam diferentes técnicas para automatizar a captura e/ou análise de imagens do animal (Akerberg *et al.*, 2019; Westhoff *et al.*, 2020; Yang *et al.*, 2018; Zhang *et al.*, 2021).

### **1.1.2 Uso do HCS e sua implementação no modelo *zebrafish***

Uma das técnicas de imageamento que tem se destacado no estudo de embriões/larvas de *zebrafish* é a análise de alto conteúdo (HCA ou HCS), que utiliza um microscópio comercial automatizado para adquirir e analisar imagens de fluorescência (ou luz transmitida) de microplacas. A Análise de Alto Conteúdo é uma abordagem que combina alta precisão e eficiência, permitindo a avaliação sistemática de atividades biológicas em larga escala. Esse método é amplamente utilizado para investigar respostas celulares ou de organismos inteiros a compostos químicos. A caracterização dessas respostas pode ser realizada por meio da aplicação de sondas fluorescentes ou utilizando organismos geneticamente modificados que expressam proteínas fluorescentes, que facilitam a quantificação de parâmetros específicos. Outra alternativa é capturar imagens de campo claro que, embora não forneçam detalhes como as de fluorescência, também podem ser úteis para extrair parâmetros gerais associados a morfologia e comportamento do objeto de estudo. Este tipo de análise não requer organismos transgênicos fluorescentes, nem a marcação de indivíduo por sondas, que podem ter efeitos tóxicos (Buchser *et al.*, 2012).

Na HCA, a análise das imagens também é feita de forma automatizada num *software* dedicado e consiste geralmente na segmentação dos objetos de interesse na imagem

(ex:células ou pequenos organismos) seguida da extração e quantificação de centenas a milhares de parâmetros morfológicos relacionados a forma, o tamanho, a intensidade e a textura do objeto. Tais parâmetros podem então ser usados isoladamente ou em conjunto para mensurar o efeito de um composto teste ou associá-lo a um fenótipo específico (Buchser *et al.*, 2012; Yozzo *et al.*, 2013). Essa última etapa pode ser aprimorada com o uso de classificadores baseados em modelos de inteligência artificial (IA) (Bray *et al.*, 2017). Alguns desses modelos também são capazes de classificar o fenótipo diretamente a partir da imagem (Jeanray *et al.*, 2015; Shang; Lin; Cong, 2020).

A segmentação dos embriões de zebrafish nas imagens tem sido realizada através de, pelo menos, duas abordagens principais. A primeira consiste na aplicação de uma sequência de filtros que visam isolar o embrião, facilitando sua identificação (Jeanray *et al.*, 2015). Alternativamente, a segmentação também tem sido realizada diretamente sobre a imagem original por meio da utilização de modelos de IA, como as redes neurais convolucionais (Teixidó *et al.*, 2018).

Os fenótipos identificados podem ser descritos com base em um ou mais parâmetros quantificáveis do objeto. Estudos anteriores demonstraram que malformação cardíaca induzidas por compostos químicos podem ser quantificadas por meio de parâmetros relacionados a área do objeto, especialmente ao se observar o aumento da região pericárdica (Yozzo *et al.*, 2013). Além disso, sinais de letalidade também puderam ser correlacionados com alterações na área corporal total dos embriões (Deal *et al.*, 2016). Outro parâmetro morfológico, conhecido como “form factor”, por sua vez, também tem se mostrado eficaz na detecção de alterações morfológicas associadas ao grau de circularidade do embrião. Esse parâmetro normalmente se encontra mais elevado em embriões que não eclodiram (Beckman, 2018).

Embora a maioria dos ensaios de HCA seja realizado com zebrafish transgênicos, que expressam proteínas fluorescentes, também é possível extrair parâmetros fenótipos (ex: morfologia do corpo, área pericárdica, fluxo sanguíneo) a partir de imagens de campo claro e usá-los para quantificar o efeito induzido por compostos ou outros agentes (Beckman, 2018; Deal *et al.*, 2016; Yozzo *et al.*, 2013). A versatilidade dos ensaios de HCA pode ser exemplificada pelos diferentes estudos com *zebrafish* que avaliam nefrotoxicidade, degeneração de neurônios, função cardiovascular, angiogênese, crescimento e proliferação de células de câncer e alterações morfológicas gerais (Deal *et al.*, 2016; Liu *et al.*, 2016; Westhoff *et al.*, 2020; Yozzo *et al.*, 2013; Zhang *et al.*, 2014).

Ao longo dos anos nosso grupo (Laboratório de Bioquímica Experimental e Computacional de Fármacos, IOC/FIOCRUZ) tem avaliado o potencial farmacológico de compostos químicos de origem comercial e de colaboração empregando diferentes técnicas experimentais e computacionais. Algumas dessas moléculas já se mostraram promissoras em modelos *in vitro* e *ex vivo* de doenças infecciosas e crônicas humanas, como a esquistossomose, a diabetes mellitus, a COVID-19 e câncer (Azevedo *et al.*, 2024; Gomes *et al.*, 2023; Saramago *et al.*, 2023; Senger *et al.*, 2024). Atualmente a seletividade destes compostos tem sido avaliada utilizando ensaios com linhagens de células humanas que, embora sejam úteis, não consideram a complexidade de um ser vivo vertebrado. Sendo assim, o uso de embriões de *zebrafish* oferece uma visão mais realista dos seus possíveis efeitos tóxicos, o que nos permite determinar, com maior antecedência, se tais moléculas merecem ou não avançar no *pipeline* de descoberta de fármacos.

Este projeto tem como objetivo o desenvolvimento de um ensaio de microscopia automatizada de alto conteúdo (HCA) para avaliar toxicidade de compostos químicos em embriões de *zebrafish*. A proposta inclui a realização de ensaios com embriões expostos ou não a substâncias comerciais conhecidas por seu potencial tóxico, visando a caracterização de alterações fenotípicas induzidas por esses compostos. Para isso, será realizado o treinamento de modelos de rede neurais convolucionais para a segmentação automática do corpo dos embriões, assim como de alguns órgãos (área pericárdica e saco vitelino), nas imagens obtidas por HCA. Também será implementado um *pipeline* de análise de imagens no *software* CellProfiler, com o objetivo de extrair e quantificar parâmetros associados a aspectos morfológicos e tentar correlacioná-los com os fenótipos observados visualmente. Por fim, será treinado um modelo de aprendizado de máquina supervisionado para a classificação fenotípica automatizada do *zebrafish*. Todas as etapas de análise serão feitas em *softwares* gratuitos e de código aberto, possibilitando que qualquer grupo de pesquisa possa implementá-lo. Inicialmente os experimentos serão realizados com uma linhagem *wild type* (selvagem/*outbred*) do peixe, uma vez que os organismos transgênicos ainda não estão disponíveis. Uma vez estabelecido, este ensaio poderá ser aplicado em diferentes projetos de pesquisa e desenvolvimento tecnológico para avaliar toxicidade. E servirá de ponto de partida para a implementação de novos ensaios baseados em HCA com *zebrafish*.

## **2 METODOLOGIA**

### **2.1 Compostos químicos**

Neste trabalho foram testados dois compostos químicos comerciais reconhecidamente tóxicos sobre o *zebrafish* (Horie *et al.*, 2017; Jeanray *et al.*, 2015): 3,4-dicloroanilina (Merck, código: 437778) e ZnSO<sub>4</sub> (Merck, código: 221376). O reagente DMSO (Merck, código: D2650) também foi avaliado por ser um solvente normalmente utilizado na solubilização de compostos químicos (Hoyberghs *et al.*, 2021). O anestésico triclaína (Merck, código: E10521) foi usado para garantir que os animais permaneçam imóveis durante a aquisição das imagens no microscópio. Tal sedativo é comumente utilizado em experimentos com *zebrafish* (Leyden *et al.*, 2022).

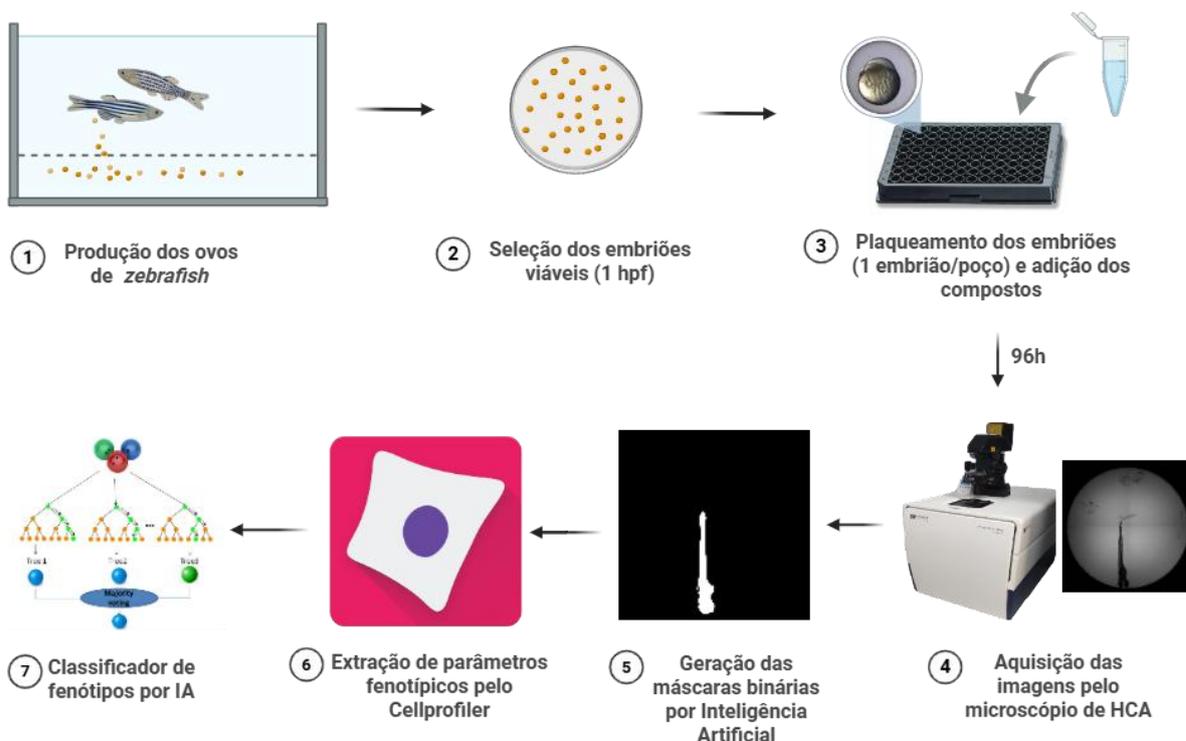
### **2.2 Zebrafish**

Os peixes *zebrafish* (*Danio rerio*) *wild type* foram obtidos e mantidos no Laboratório de Fisiologia (Instituto Nacional de Controle de Qualidade em Saúde – INCQS/Fiocruz, licenciado pela CEUA/Fiocruz nº LW-38/19), sob a chefia da Dra. Renata Jurema Medeiros, colaboradora deste projeto. Para manutenção e criação dos peixes foram utilizadas as condições descritas na literatura (Westerfield, 2007). Os animais foram mantidos a uma temperatura de 26°C ao longo de todos os experimentos. No final de cada experimento, os peixes foram anestesiados por 10 minutos com triclaína 0,02% e submetidos à eutanásia utilizando o método de imersão em água com gelo durante 20 minutos e submersão em solução de hipoclorito de sódio (500 mg/L) de acordo com a Resolução Normativa CONCEA nº 37/2018.

### **2.3 Fluxo de trabalho proposto para o ensaio de toxicidade baseado em HCA**

Uma vez padronizado, o ensaio de HCA proposto neste projeto seguirá o fluxo de trabalho ilustrado na Figura 3. O estado atual de desenvolvimento de cada etapa será descrito em mais detalhes nos próximos tópicos.

Figura 3 - Fluxo de trabalho proposto para o ensaio. hpf – horas após a fertilização; HCA – Análise de Alto Conteúdo.



Fonte: Criado por Biorender

## 2.4 Ensaio de toxicidade por microscopia de alto conteúdo

### 2.4.1 Preparação dos ensaios com embriões de *zebrafish*

O ensaio de toxicidade sobre embriões de *zebrafish* realizou-se tendo como base o teste 236 da OECD (OECD, 2013). Resumidamente, após a produção dos ovos, os embriões de *zebrafish* (1 hpf) foram coletados e adicionados a uma placa de Petri preenchida com meio embriônico, também conhecido como meio E3 (NaCl 14,69 g/L; KCl 0,63 g/L; CaCl<sub>2</sub>-2H<sub>2</sub>O 2,43 g/L; MgSO<sub>4</sub>-7H<sub>2</sub>O 4,07 g/L diluído em água miliQ). Os embriões considerados visualmente como viáveis (i.e. que não possuíam coágulo no interior) foram transferidos para placas de Petri contendo concentrações diferentes dos compostos teste. Em seguida, volumes de 250 µL contendo 1 embrião foram coletados das placas de Petri e adicionados a cada poço de uma microplaca de 96 poços preta e com o fundo transparente (Corning Incorporated, código: 3603). Os peixes foram distribuídos apenas nos poços centrais das microplacas (60 poços) enquanto os demais poços foram preenchidos com meio E3. Isto foi realizado para reduzir o efeito de evaporação que pode ocorrer ao longo do tempo.

Foram realizadas 12 replicadas para cada uma das seguintes condições: embriões mantidos apenas com meio E3 (controle negativo) ou tratados com composto teste (DMSO, 3,4-dicloroanilina ou ZnSO<sub>4</sub>). Os compostos foram testados em concentrações variáveis para avaliar o efeito de cada concentração na resposta biológica: DMSO (0,125 – 1%), 3,4-dicloroanilina (0,5 – 4 mg/L) e ZnSO<sub>4</sub> (0,005 – 300 mg/L).

#### **2.4.2 Aquisição das imagens por HCA**

Após 96h de incubação com os compostos testes, as microplacas contendo os embriões de *zebrafish* foram levados a Plataforma de Bioensaios e Triagem de Fármacos (IOC/FIOCRUZ) para a captura das imagens. Primeiramente os *zebrafish* foram anestesiados com tricaina 0,02% e centrifugados a 300 rpm a 26°C por 4 minutos. Em seguida, imagens de campo claro das microplacas foram adquiridas no sistema de HCA ImageXpress Confocal (Molecular Devices) com uma objetiva de 4X e usando um binning de 1. Para cada poço da microplaca foram capturadas quatro imagens: duas dos quadrantes superiores e duas dos quadrantes inferiores do poço.

#### **2.4.3 Análise das imagens**

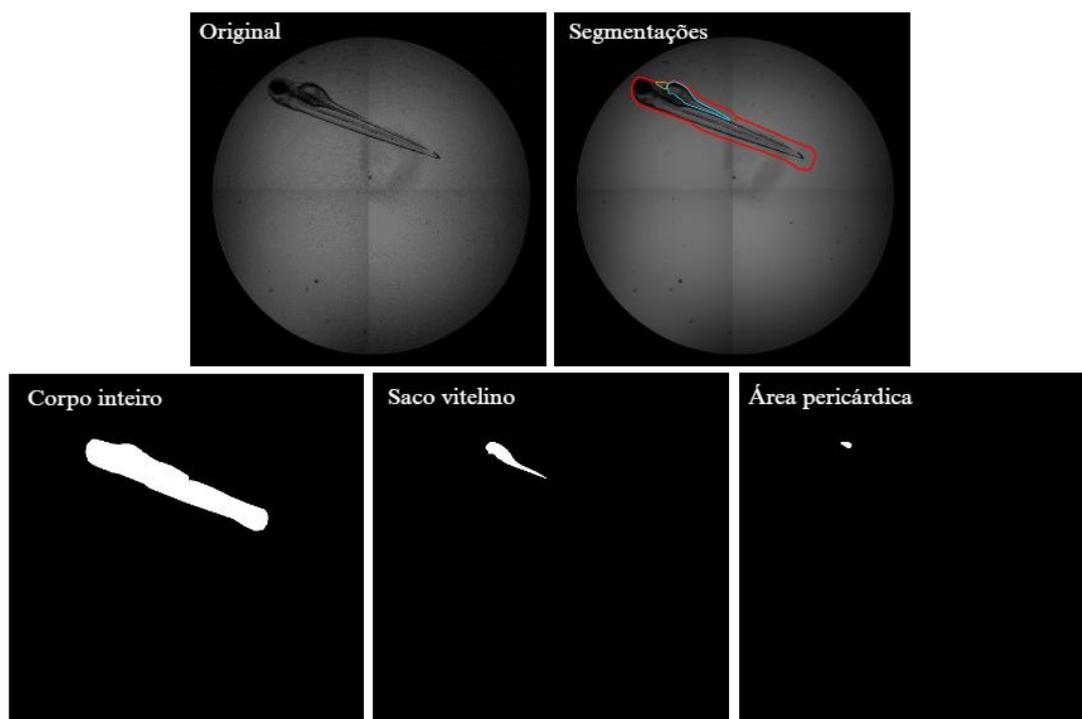
Inicialmente as imagens dos embriões do *zebrafish* foram concatenadas no programa CellProfiler versão 4.2.1 (Lamprecht; Sabatini; Carpenter, 2007) para gerar uma montagem contendo a imagem completa de cada poço da microplaca. Em seguida, os *zebrafish*, assim como seus órgãos (área pericárdica e saco vitelino), foram segmentados nas imagens de forma semiautomática (i.e. segmentação visual auxiliada por computador) para a geração de dados (máscaras binárias) que foram usados no treinamento de modelos de IA (subtópico 2.4.3.1). Tais modelos serão incorporados ao ensaio futuramente, permitindo que a segmentação seja realizada de forma totalmente automática.

##### **2.4.3.1 Treinamento de modelos de inteligência artificial para segmentação automática dos embriões de *zebrafish* e de seus órgãos**

Antes de realizar o treinamento dos modelos foi necessário selecionar um conjunto de imagens e suas respectivas máscaras binárias (i.e. versões preto/branco das imagens originais nas quais os pixels brancos representam o peixe (ou seus órgãos) e os pretos todo o resto da imagem) (Figura 4). Tais máscaras foram geradas em duas etapas. A primeira consistiu na anotação semiautomática das imagens utilizando o programa Anylabeling

(<https://anylabeling.nrl.ai/>) auxiliado pela rede neural SAM (<https://segment-anything.com/>). Ao final desta etapa foram gerados arquivos contendo as coordenadas do contorno de cada objeto (corpo inteiro, saco vitelino ou área pericárdica do *zebrafish*) na imagem. A partir destes dados foi possível gerar as máscaras binárias através de um *script* escrito na linguagem de programação *python*.

Figura 4 - Imagem original e máscaras binárias do corpo inteiro, saco vitelino e área pericárdica. Os contornos coloridos na imagem superior à direita representam as regiões segmentadas do corpo (vermelho), saco vitelino (azul) e área pericárdica (verde).



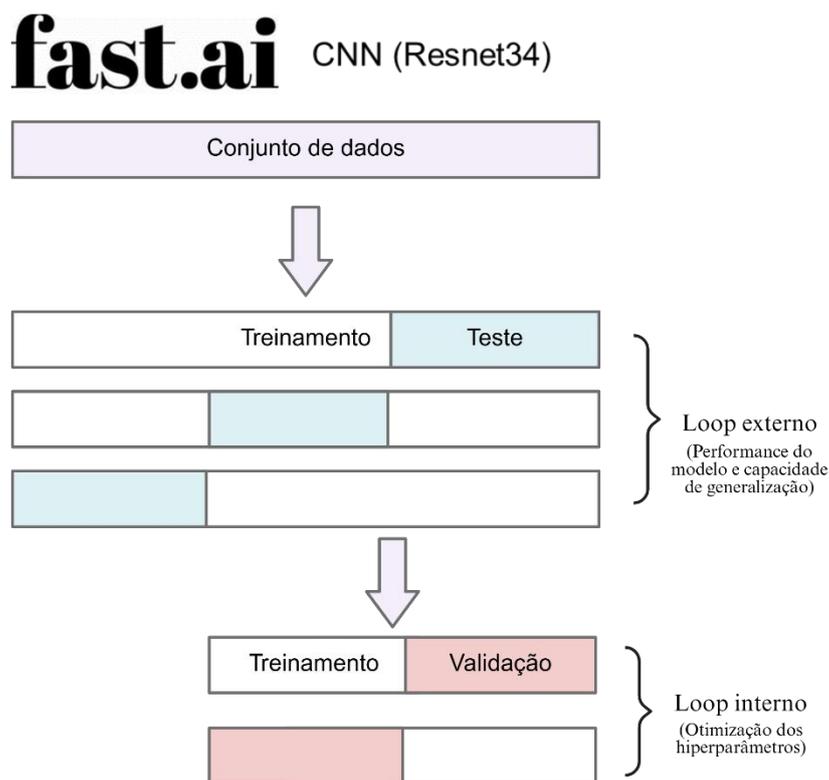
Fonte: LABECFAR, 2025

Os conjuntos de imagens/máscaras foram utilizados no treinamento de um modelo de rede neural convolucional (Resnet 34) pré-treinado com o banco de imagens ImageNet (<https://www.image-net.org/>), disponível na biblioteca fast.ai (<https://www.fast.ai/>) para python. Diferentes instâncias do modelo foram usadas para treinar a segmentação do corpo, saco vitelino e área pericárdica isoladamente. A validação dos modelos foi realizada pela técnica de validação cruzada aninhada utilizando funções da biblioteca do scikit-learn (<https://scikit-learn.org/stable/>) conforme ilustrado na Figura 5. Esta técnica permite ao mesmo tempo otimizar os hiperparâmetros (i.e. parâmetros que não são ajustados durante o treinamento) do modelo e avaliar seu desempenho de forma imparcial (Lobato *et al.*, 2024). Para a otimização dos hiperparâmetros foi utilizada a biblioteca do Optuna

(<https://optuna.org/>) para python. No total, quatro parâmetros foram otimizados em 100 tentativas (Tabela A1). Neste processo foi utilizado o algoritmo de *prunning* do optuna para interromper precocemente a otimização dos hiperparâmetros quando o modelo não demonstrava uma melhora ao decorrer do treinamento.

A performance do modelo ao longo de todo o treinamento foi calculada por meio da métrica Dice, comumente usada na avaliação de modelos de segmentação aplicados a imagens biomédicas (Taha; Hanbury, 2015). Essa métrica avalia o grau de sobreposição entre as máscaras preditas pelo modelo e as máscaras de referência, sendo que o valor 1 indica sobreposição total, ou seja, concordância perfeita entre as segmentações. Ao final da validação cruzada aninhada, o valor de Dice foi determinado tanto sobre o conjunto de treinamento quanto de teste em cada uma das três das versões do modelo (ou seja, os modelos treinados em cada um dos três loops externos). O desempenho final do modelo foi então determinado pela média dos valores de Dice obtidos nos conjuntos de teste. Além de avaliar a performance, também geramos as curvas de perda de cada versão do modelo para avaliar seu comportamento ao longo do treinamento.

Figura 5 - Esquema geral de validação cruzada aninhada com 3 loops externos e 2 internos. Neste tipo de validação, o conjunto total de dados é inicialmente dividido em conjunto de treinamento e teste de três formas diferentes (3 loops externos). Por sua vez, o conjunto de treinamento de cada loop externo é dividido num conjunto interno de treinamento e outro de validação de duas formas diferentes (2 loops internos). No loop interno foi realizada a otimização dos hiperparâmetros do modelo via optuna. Os valores dos hiperparâmetros selecionados após a otimização foram aplicados no treinamento dos modelos com o conjunto de treinamento do loop externo e depois avaliados no conjunto teste correspondente. Essa abordagem permitiu avaliar o desempenho do modelo, refletindo sua capacidade de generalização - ou seja, sua habilidade de realizar previsões precisas em dados não utilizados durante o treinamento.



Fonte: Adaptado de Lobato *et al* 2024

### 2.4.3.2 Extração automática de parâmetros fenotípicos do *zebrafish*

Após a etapa de segmentação, as imagens originais e suas respectivas máscaras binárias (geradas de forma semiautomática) foram analisadas num *pipeline* customizado do programa CellProfiler v.4.2.6 (Lamprecht; Sabatini; Carpenter, 2007). A partir deste *pipeline* foi possível quantificar, de cada *zebrafish* (corpo, área pericárdica e saco vitelino), cerca de 744 parâmetros relacionados à sua forma, textura, granulosidade e intensidade de *pixels*.

### 2.4.3.3 Análise dos fenótipos

Em virtude do reduzido número de imagens e da variedade de alterações morfológicas encontradas, os embriões foram classificados visualmente em três fenótipos gerais: “normal” (embriões considerados normais de acordo com os critérios da OECD), “não eclodido” e “alterado” (embriões com quaisquer alterações visíveis).

Em seguida, investigamos possíveis correlações entre parâmetros extraídos por meio do CellProfiler e os fenótipos previamente definidos. Até o momento, esta análise foi realizada apenas com os parâmetros “area” e “form factor” do módulo MeasureObjectSizeShape, uma vez que tais descritores já foram associados com fenótipos induzidos com compostos químicos (Yozzo *et al.*, 2013; Deal *et al.*, 2016). O primeiro quantifica o número de pixels do objeto e o segundo é calculado pela fórmula  $4*\pi*\text{Área}/\text{Perímetro}^2$ . O valor do “form factor” igual a 1 descreve um objeto perfeitamente circular, esse parâmetro foi utilizado exclusivamente na análise do corpo inteiro do embrião. As áreas pericárdicas e do saco vitelino foram posteriormente normalizadas pela área corpórea do embrião. Os gráficos e as análises estatísticas usadas para a comparação entre os fenótipos foram realizados no programa GraphPad Prism version 8 para Windows, GraphPad Software, Boston, Massachusetts USA, [www.graphpad.com](http://www.graphpad.com).

#### **2.4.3.4 Classificador baseado em IA**

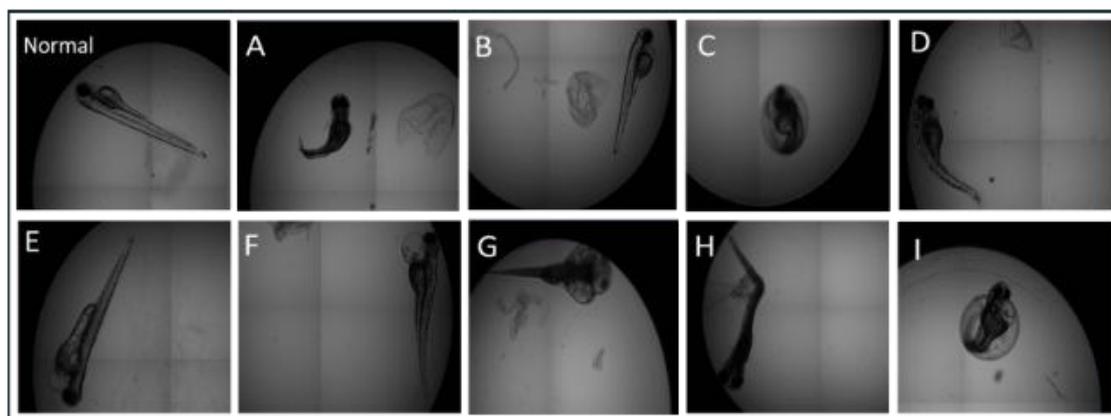
Em uma etapa subsequente, os 744 parâmetros extraídos por meio do *software* CellProfiler, foram utilizados como entrada para treinamento de um modelo de classificação fenotípica baseado no algoritmo de aprendizado de máquina XGBoost, implementado na linguagem de programação *python* ([https://xgboost.readthedocs.io/en/release\\_3.0.0/](https://xgboost.readthedocs.io/en/release_3.0.0/)). Assim como na etapa da segmentação, empregou-se uma validação cruzada aninhada (3 loops externos e 2 internos) (Figura 5) e também foi implementada a biblioteca do Optuna para a otimização dos hiperparâmetros utilizando 1000 tentativas (Tabela A2). A métrica utilizada para avaliar a performance do modelo durante o treinamento foi o *Matthews correlation coefficient* (MCC). Diferente de outras métricas, como acurácia e precisão, o MCC considera de forma equilibrada todas as quatro categorias da matriz de confusão: verdadeiros positivos (TP), falsos positivos (FP), falsos negativos (FN) e verdadeiros negativos (TN), proporcionando uma medida mais robusta e informativa da qualidade do classificador (Chicco; Jurman, 2020). Ao final da validação cruzada aninhada, o desempenho final do modelo foi calculado como a média dos valores de MCC obtidos por suas três versões (uma para cada um dos folds externos) em seus respectivos conjuntos de teste. Além do MCC, as métricas precisão, sensibilidade, F1, acurácia balanceada e área sob a curva ROC (AUC), também foram determinadas para avaliar a performance do modelo.

### 3 RESULTADOS

#### 3.1 Análises dos fenótipos do *zebrafish*

Durante os experimentos foi identificada uma variedade de fenótipos dos embriões de *zebrafish* tratados ou não com os compostos teste. Os fenótipos que apareceram com maior frequência estão exemplificados na Figura 6. Além do fenótipo associado ao peixe normal, observamos 9 variações fenotípicas (A-I). Nelas é possível notar uma série de alterações morfológicas, como edema pericárdico (ex: **F**) e do saco vitelino (ex: **E**), curvatura anormal da coluna (ex: **D**) e não eclosão do ovo (ex: **C**).

Figura 6 - Principais fenótipos observados em embriões de *zebrafish* tratados ou não com os compostos teste. Além do fenótipo associado com o peixe normal, outros fenótipos com alterações morfológicas foram encontrados (A-I).



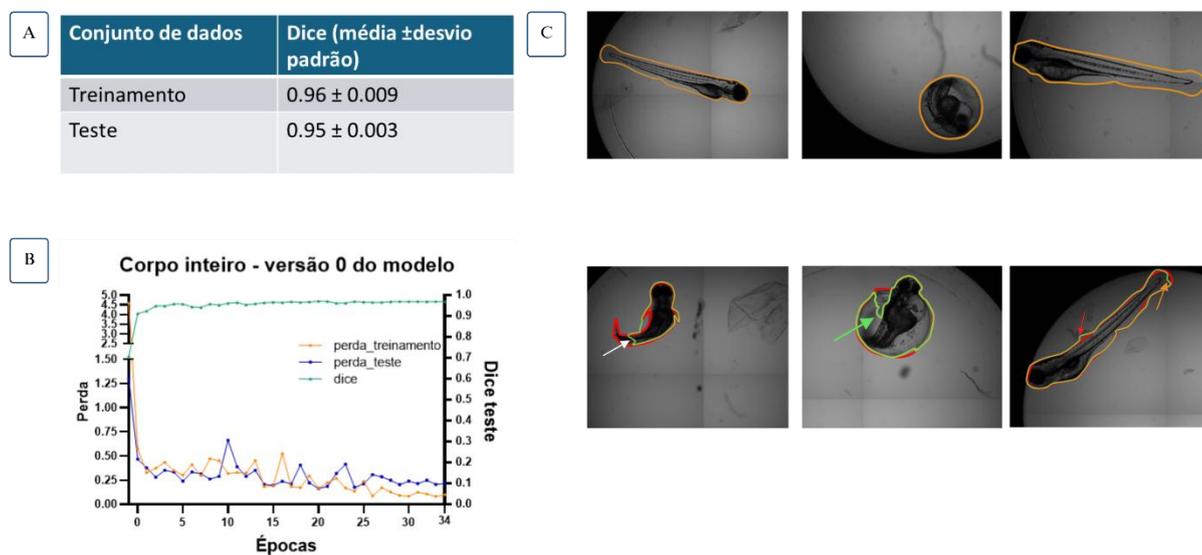
Apesar da variedade de fenótipos alterados encontrados, não foi possível determinar se tais alterações foram induzidas ou não pelos compostos testados, uma vez que todos os experimentos realizados foram reprovados em um ou mais critérios mínimos de validação de ensaio sugeridos para o teste 236 da OECD (OECD, 2013). Na maioria dos casos, os ensaios foram invalidados devido os controles negativos (i.e. peixes expostos apenas ao meio E3) apresentarem uma quantidade significativa de fenótipos alterados (ex: número de ovos eclodidos menor que 80 %, ausência de somitos) e os controles positivos (peixes tratados com 4,0 mg/L 3,4-dicloroanilina) não provocarem uma mortalidade mínima de 30%.

As imagens obtidas nesses experimentos, no entanto, foram usadas em três etapas importantes da padronização do ensaio: treinamento de um modelo de IA de segmentação dos embriões (corpo inteiro, saco vitelino e área pericárdica) nas imagens, quantificação automática de seus parâmetros morfológicos e treinamento de um classificador fenotípico baseado em IA. Os resultados de cada etapa estão descritos a seguir.

### 3.2 Segmentação das imagens por IA

Após o treinamento, o modelo de rede neural convolucional usado para segmentação do corpo inteiro do embrião apresentou um valor de Dice de  $0,96 \pm 0,009$  para o conjunto de treinamento e  $0,95 \pm 0,003$  para o conjunto teste de imagens (Figura 7A). É possível observar no gráfico de perda da versão do modelo de melhor performance (Figura 7B) que o perfil de queda no valor de perda ao longo das épocas foi semelhante entre o conjunto de treinamento e teste. Já o Dice para o conjunto teste aumentou após as épocas iniciais se mantendo praticamente inalterado até o final do treinamento. O mesmo ocorreu com as outras duas versões do modelo (Figura A1). Na parte superior da Figura 7C, são exemplificados casos nos quais houve um alto grau de sobreposição entre as máscaras previstas pelo modelo de IA e as geradas de forma semiautomática, enquanto na parte inferior temos exemplos de imagens que apresentaram sobreposição parcial.

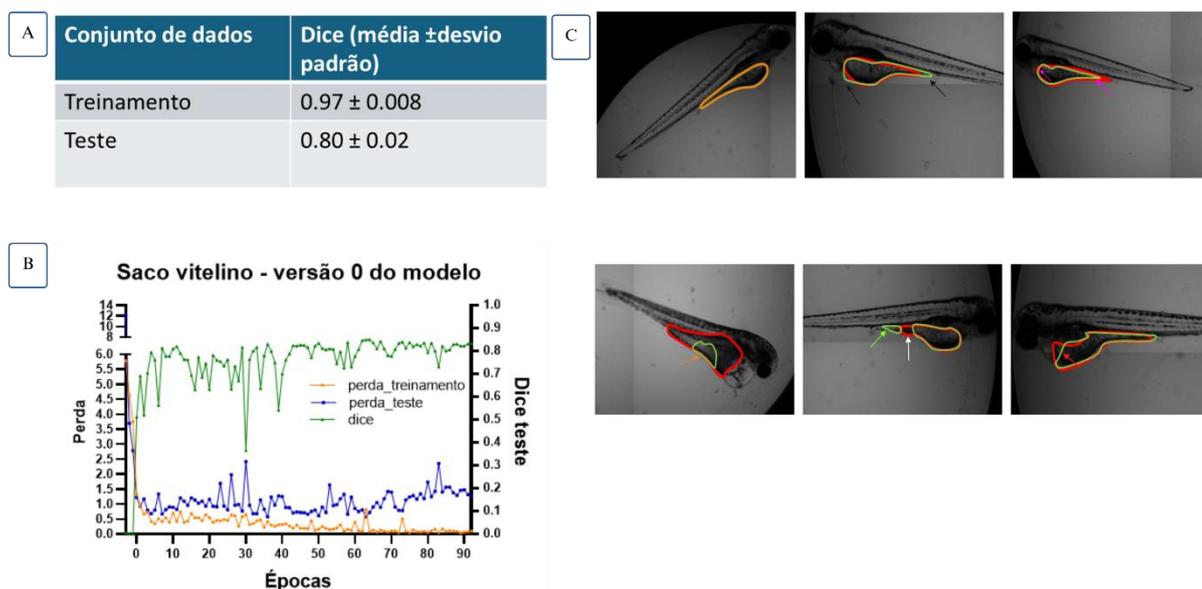
Figura 7 - Resultados obtidos após o treinamento do modelo de segmentação para o corpo inteiro de embrião de *zebrafish*. (A) Valores de Dice calculados para o conjunto de treinamento e teste. (B) Curva de perda gerada durante o treinamento da versão do modelo com maior valor de Dice. (C) Comparação entre as máscaras previstas pela rede neural (contornos em verde) e as geradas de forma semiautomática (contorno em vermelho). As regiões de sobreposição entre as máscaras estão representadas em contorno laranja. As setas apontam para regiões onde não houve sobreposição.



Na Figura 8 são mostrados os resultados obtidos com o modelo de segmentação do saco vitelino. O valor médio de Dice para o conjunto de treinamento foi de  $0.97 \pm 0.008$  enquanto para o conjunto teste  $0.80 \pm 0.02$  (Figura 8A). Na curva de aprendizado do modelo de melhor performance (Figura 8B), é possível observar um declínio contínuo no valor de perda

associado ao conjunto de treinamento, enquanto a do conjunto teste se mantém estável ao longo da maior parte do treinamento, aumentando progressivamente após 70 épocas. Já o valor de Dice aumentou após as primeiras épocas oscilando em torno de 0.8 até o final do treinamento. Na Figura 8C as imagens superiores mostram o alto grau de sobreposição entre as máscaras semiautomáticas e as previstas pelo modelo enquanto as inferiores mostram casos de sobreposição parcial.

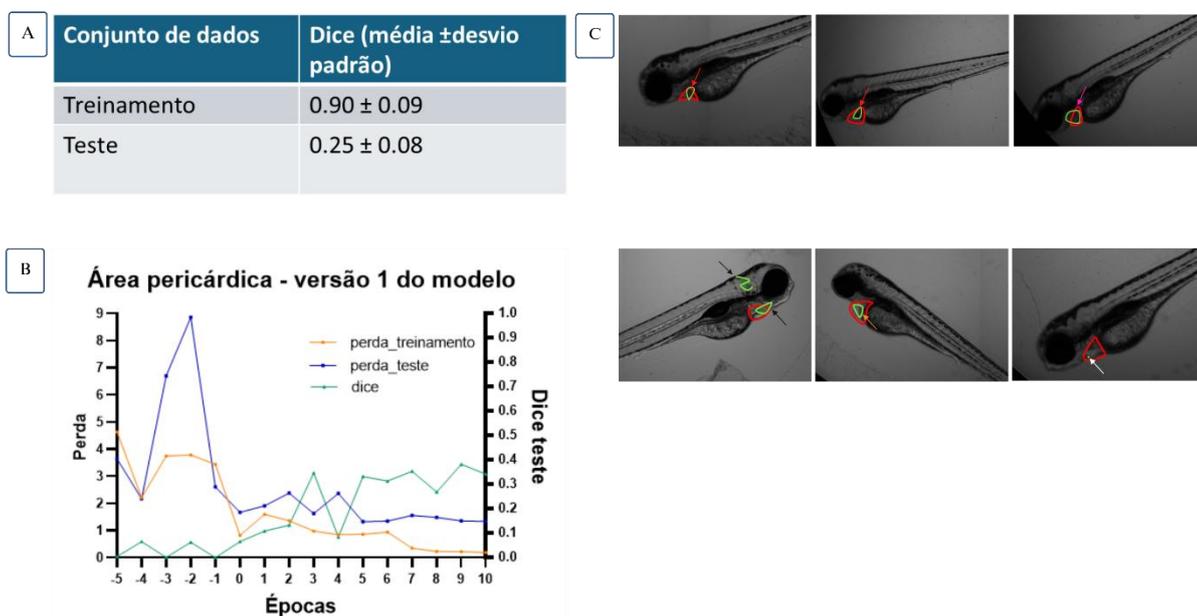
Figura 8 - Resultados obtidos após o treinamento do modelo de segmentação para o saco vitelino do embrião de *zebrafish*. (A) Valores de Dice calculados para o conjunto de treinamento e teste. (B) Curva de perda gerada durante o treinamento da versão do modelo com maior valor de Dice. (C) Comparação entre as máscaras previstas pela rede neural (contornos em verde) e as geradas de forma semiautomática (contorno em vermelho). As regiões de sobreposição entre as máscaras estão representadas em contorno laranja. As setas apontam para regiões onde não houve sobreposição.



A segmentação da área pericárdica apresentou os menores valores de Dice tanto para o conjunto de treinamento ( $0.90 \pm 0.09$ ) quanto para o teste ( $0.25 \pm 0.08$ ) (Figura 9A). A curva de perda do conjunto de treinamento continuou caindo ao longo do treinamento, enquanto a do teste parece ter estabilizado no final do treinamento (Figura 9B). Todas as imagens analisadas do conjunto teste apresentaram baixa sobreposição entre as máscaras geradas de forma semiautomática e as previstas pelos modelos (Figura 9C).

Figura 9 - Resultados obtidos após o treinamento do modelo de segmentação para a área pericárdica do embrião de *zebrafish*. (A) Valores de Dice calculados para o conjunto de treinamento e teste. (B) Curva de perda gerada durante o treinamento da versão do modelo com maior valor de Dice. (C) Comparação entre as máscaras previstas pela rede neural (contornos em verde) e as geradas de forma semiautomática (contorno em vermelho).

e as geradas de forma semiautomática (contorno em vermelho). As regiões de sobreposição entre as máscaras estão representadas em contorno laranja. As setas apontam para regiões onde houveram sobreposições parciais.

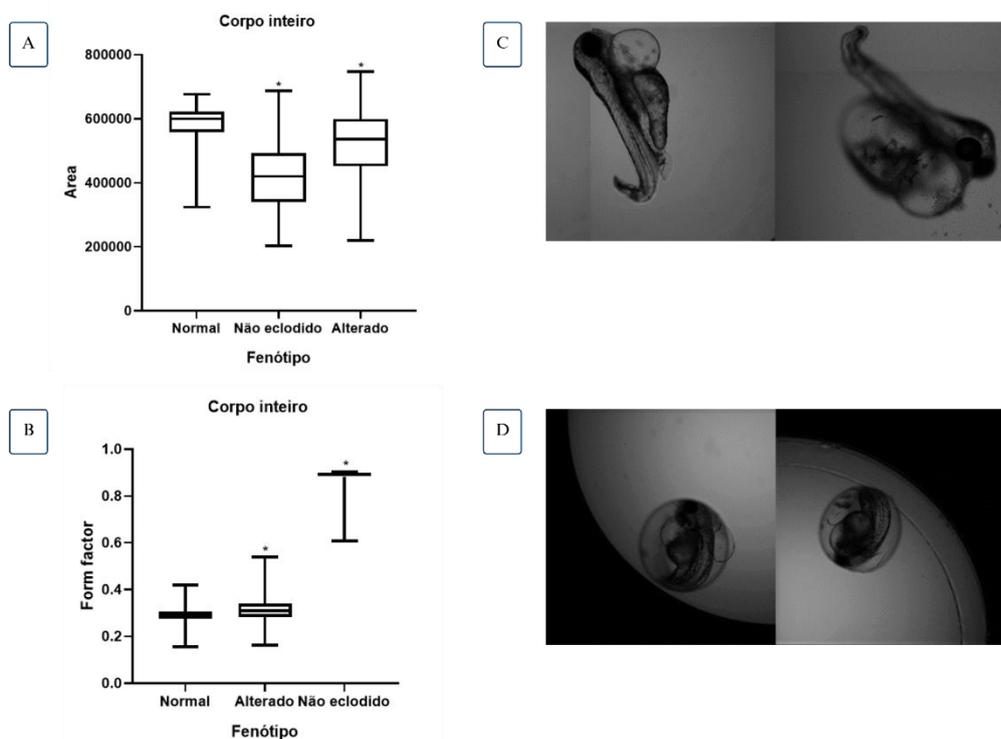


### 3.3 Avaliação do efeito de compostos teste sobre embriões de zebrafish

Posteriormente classificamos os embriões em três fenótipos gerais (“normal”, “não eclodido” e “alterado”) e tentamos relacionar tais fenótipos com os parâmetros quantificados pelo CellProfiler. Até o momento, foram avaliados apenas os parâmetros “area” (corpo inteiro, área pericárdica e saco vitelino) e “form fator” (corpo inteiro).

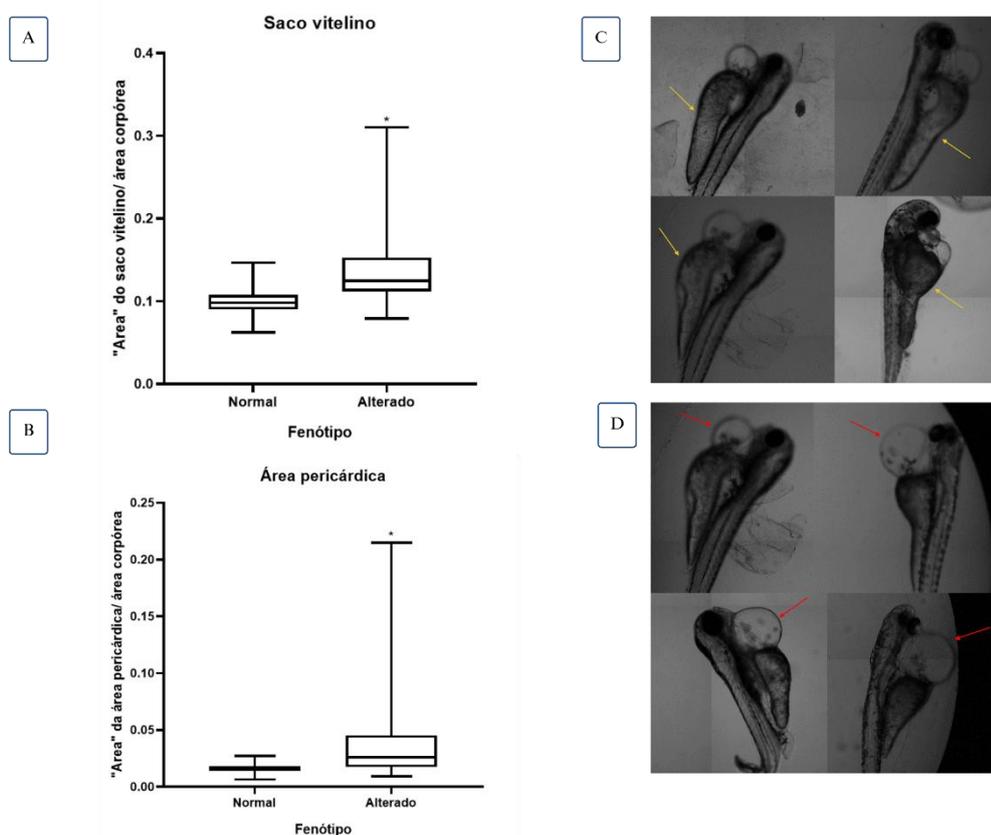
A Figura 10 apresenta os gráficos referentes aos parâmetros de “area” e “form factor” quantificados a partir das imagens do corpo inteiro dos embriões para cada grupo fenotípico analisado. No parâmetro “area”, o fenótipo “não eclodido” apresentou uma redução significativa em comparação ao grupo “normal”, resultado também observado para o grupo “alterado” (Figura 10A). Já na Figura 10B é mostrado o gráfico comparando o parâmetro “form factor” entre os fenótipos. Tanto o grupo “alterado” quanto o “não alterado” apresentaram valores significativamente maiores que os embriões do grupo “normal”. O grupo que mais se destacou foi o “não eclodido”, por apresentar valores significativamente maiores para este parâmetro. Na Figura 10C-D são mostrados exemplos de imagens de embriões com valores mais elevados de “form factor” nos grupos “alterado” e “não eclodido”, respectivamente.

Figura 10 - Avaliação dos parâmetros “area” (A) e “form fator” (B) nos diferentes fenótipos de *zebrafish*. Em (C-D) são exemplificadas imagens de embriões classificados como “alterado” (C) e “não eclodido” (D) que apresentaram valores mais elevados do parâmetro “form fator” para cada grupo. \* Diferença estatísticas ( $p < 0,05$ ) em relação ao fenótipo normal segundo o teste ANOVA, seguido do pós-teste de Dunnett.



O parâmetro “area” também foi empregado na avaliação dos fenótipos com foco em estruturas específicas dos embriões: área pericárdica e saco vitelino (Figura 11). Os indivíduos classificados como “alterado” apresentaram valores significativamente mais elevados de ambos os parâmetros (Figura 11A-B). Nas Figuras 11C e D são mostrados exemplos de embriões do fenótipo “alterado” que apresentaram valores de área de saco vitelino e área pericárdica, respectivamente.

Figura 11 - Avaliação da área do saco vitelino (A) e da área pericárdica (B) em embriões de *zebrafish* com fenótipos "normal" e "alterado". Em (C) e (D) exemplos de embriões do fenótipo "alterado" com valores elevados de área de saco vitelino (C) e área pericárdica (D), respectivamente. As setas em amarelo mostram o edema no saco vitelino e as em vermelho apontam o edema pericárdico. \* diferença estatística ( $p < 0,05$ ) segundo teste t.



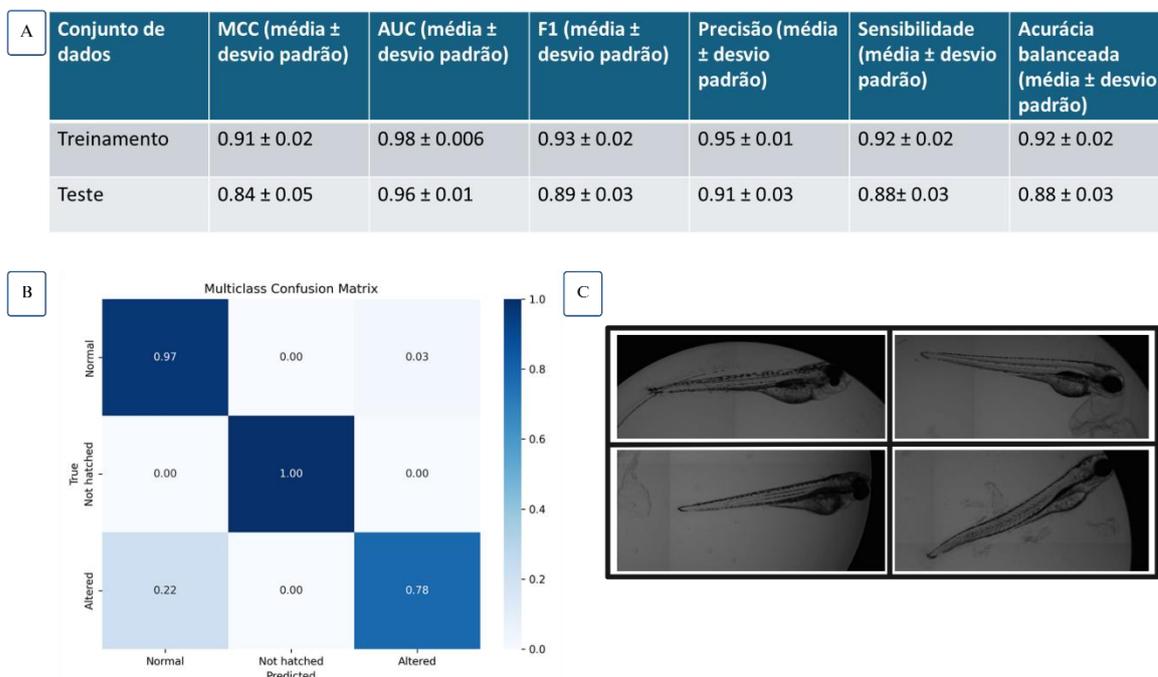
### 3.4 Modelo de classificação gerado por IA

Além do estudo feito com os parâmetros “área” e “form factor” de forma isolada também realizamos uma análise multiparamétrica dos dados. Para isso, foi treinado um modelo de aprendizado de máquina supervisionado, conhecido como XGBoost, com o objetivo de construir um classificador para categorização dos fenótipos “normal”, “alterado” e “não eclodido” a partir da quantificação de todos os parâmetros extraídos pelo CellProfiler. Os resultados obtidos após o treinamento do modelo estão mostrados na Figura 12 e na Figura A4.

De acordo com a Figura 12A, a melhor versão do classificador apresentou um valor de MCC de  $0.91 \pm 0.02$  para o conjunto de treinamento e  $0.84 \pm 0.05$  para o teste. As demais métricas de performance ficaram em torno de 0.9 para o conjunto teste e um pouco mais elevadas para o conjunto de treinamento. Na matriz de confusão gerada após a avaliação sobre

o conjunto teste é possível notar que o modelo identificou corretamente 100% das imagens do grupo “não eclodido”, 97% do grupo “normal” e 78% do “alterado” (Figura 12B). As imagens da Figura 12C representam algumas imagens de embriões da classe “alterado” que o modelo de IA classificou erroneamente como “normal”.

Figura 12 – Performance do classificador baseado em IA. (A) Valores de média e desvio padrão do MCC e de outras métricas de performance para os conjuntos de treinamento e teste; (B) Resultado da matriz de confusão da versão do modelo XGBoost que apresentou o maior valor de MCC; (C) Exemplos de embriões com fenótipo “alterado” que foram classificados como “normal”.



## 4 DISCUSSÃO

Embora os ensaios de toxicidade com os embriões de zebrafish tenham sido invalidados de acordo com os critérios da OECD, as imagens adquiridas nos experimentos foram utilizadas para treinamento de dois tipos de modelos de aprendizado de máquina supervisionados: modelos de segmentação e classificador fenotípico.

Os resultados obtidos após o treinamento do modelo de segmentação do corpo inteiro do embrião foram satisfatórios uma vez que os valores de Dice tanto no conjunto de treinamento

(0,96) quanto no conjunto teste (0,95) foram muito próximos a 1, indicando que as máscaras preditas pelo modelo de IA foram muito semelhantes as máscaras geradas de forma semiautomática (Figura 7A). O valor elevado de Dice no conjunto teste significa que o modelo é capaz de manter uma boa performance mesmo quando usado na predição de dados que não foram usados no treinamento, ou seja, apresenta uma boa capacidade de generalização.

Na Figura 7B é mostrada a curva de perda para versão do modelo que apresentou o melhor desempenho (i.e., maior valor de Dice). Neste tipo de gráfico é possível acompanhar o valor de perda associado tanto ao conjunto de treinamento quanto teste ao longo do treinamento do modelo. Tal valor é calculado a partir da função de perda (*loss function*) da rede neural que quantifica o grau de discrepância entre as segmentações obtidas de forma semiautomática e aquelas geradas pelo modelo de inteligência artificial. Espera-se que, ao longo das épocas, os valores de perda diminuam, indicando que o modelo está aprendendo com os dados do treinamento. Este é o caso tanto para a curva observada para o conjunto de treinamento quanto do teste (Figura 7B), que indica que os atributos aprendidos no treinamento também generalizam bem para dados não vistos anteriormente. Comportamento similar também pode ser visto nas curvas de perda das demais versões do modelo (Figura A1). Os resultados da segmentação são exemplificados na Figura 7C onde as imagens superiores mostram casos nos quais houve uma sobreposição perfeita entre as máscaras geradas de forma semiautomática e preditas pela rede neural, representando a maior parte das imagens do grupo teste, e nas imagens inferiores algumas poucas falhas cometidas pelo modelo, como: falha na segmentação da cauda do peixe (seta branca), não identificou parte do córion (seta verde), segmentou uma pequena parte do fundo do poço (seta vermelha) e não identificou uma pequena região translúcida da cauda do embrião (seta laranja). Em conjunto, os resultados sugerem que este modelo está próximo de ser usado de forma automática no ensaio.

O modelo de segmentação para o saco vitelino apresentou um valor de Dice superior para o conjunto de treinamento (0,97) em comparação ao conjunto teste (0,80), uma diferença de cerca de 17%, sugerindo uma possível ocorrência de um fenômeno conhecido como *overfitting* (Figura 8A). Esse fenômeno acontece quando o modelo de aprendizado de máquina se ajusta excessivamente aos dados do conjunto de treinamento, o que pode comprometer a sua capacidade de generalização e, conseqüentemente, seu desempenho sobre novos conjuntos de dados. Esse efeito pode ser melhor visualizado através da curva de perda (Figura 8B). Enquanto o valor da função de perda do conjunto de treinamento apresentou uma

redução contínua ao longo de todo o treinamento – indicando um bom ajuste aos dados -, o valor correspondente no conjunto de teste apresentou uma queda inicial, mas depois estabilizou e aumentou progressivamente ao final do treinamento. Essa diferença entre os perfis das duas curvas também pode ser vista nas outras versões do modelo (Figura A2). Tal comportamento revela que o modelo passou a priorizar características específicas do conjunto de treinamento que não estão presentes no conjunto teste. Em determinados casos, como os ilustrados na Figura 8C, o modelo apresentou desempenho satisfatório, com poucas falhas localizadas (indicadas pelas setas preta e lilás). No entanto, em outras situações, exemplificadas nas imagens inferiores, o desempenho foi significativamente inferior: o modelo segmentou apenas uma pequena porção do saco vitelino (seta laranja), deixou de identificar toda a sua extensão (seta branca), interpretou erroneamente uma única estrutura como duas distintas (seta verde) e não conseguiu segmentar corretamente parte do órgão do peixe (seta vermelha). Em alguns casos, o modelo começou a identificar as regiões mais claras do órgão como a “região de interesse”.

Em conjunto, os resultados demonstraram que o modelo de segmentação do saco vitelino ainda não está pronto para ser usado no ensaio. É provável que a performance do modelo melhore após a coleta de mais imagens, o que pode também reduzir o *overfitting*. Outras estratégias, como a aplicação de métodos de regularização, *data augmentation* (i.e. aumento artificial de dados através da aplicação de filtros que alteram características da imagem, como brilho/contraste e rotação) ou até mesmo a utilização de modelos menos complexos, também podem contribuir para a obtenção de melhores resultados.

O modelo de segmentação da área pericárdica foi o que apresentou os piores resultados até o momento. O valor médio de Dice para o conjunto teste (0,25) foi baixo e muito inferior ao observado no conjunto de treinamento (0,90), uma diferença cerca de 72% (Figura 9A). Assim como ocorreu no modelo de segmentação do saco vitelino, também nota-se sinais de possível *overfitting* (Figura 9B). Este efeito fica mais evidente quando analisadas as outras duas versões do modelo (Figura A3). Como esperado, o modelo apresentou muitas falhas de segmentação (Figura 9C). Em alguns casos, como ilustrado nas imagens superiores da Figura 9C, a máscara predita se aproximou da gerada semiautomaticamente (setas vermelha e lilás). Em contrapartida, nas imagens inferiores, são observadas falhas mais significativas, incluindo a identificação incorreta do otólito e a segmentação indevida de parte da brânquia (setas pretas), além da delimitação incompleta da região de interesse (seta laranja) ou mesmo da segmentação restrita a um único pixel (seta branca).

A dificuldade na identificação da área pericárdica pode ser atribuída a diversos fatores. Primeiramente, trata-se de uma estrutura anatômica de dimensões reduzidas, o que por si só já dificulta sua delimitação precisa. Ademais, as imagens utilizadas, obtidas pelo HCA, encontram-se em escalas de tons de cinza, o que compromete a percepção de características cromáticas importantes para a diferenciação tecidual. Outro fator relevante é a necessidade de redimensionamento das imagens durante o processo de treinamento da rede neural convolucional, no qual resoluções originalmente altas (4096 x 4096 *pixels*) são reduzidas para 1024 x 1024 *pixels*, ocasionando perda de detalhes visuais relevantes para a segmentação precisa da área em questão. Essa redução foi necessária devido ao alto consumo de memória (GPU) que ocorre durante o treinamento. Sendo assim, antes de aplicar este modelo de forma automática no ensaio, será necessário primeiro ampliar o número de imagens, assim como aplicar estratégias citadas anteriormente para o modelo de segmentação do saco vitelino.

Na análise fenotípica, foi possível observar uma diferença estatisticamente significativa entre as áreas corpóreas dos embriões do grupo “alterado” e “não eclodido” em relação ao fenótipo considerado “normal” (Figura 10A), indicando que os embriões com alterações morfológicas em média são menores que os normais. No parâmetro “form factor” (Figura 10B), a classe “não eclodido” se destacou dos demais fenótipos por apresentar valores elevados (i.e. próximos de 1), o que pode ser explicado pelo formato predominantemente circular dos embriões correspondentes (Figura 10D). No entanto, a classe “alterado” também apresentou um aumento significativo nesse parâmetro em relação ao fenótipo “normal”, possivelmente em decorrência da presença de alguns embriões de menor tamanho e com alterações morfológicas que conferem aspecto circular (Figura 10C).

O parâmetro “area” também foi analisado com foco em dois órgãos específicos: saco vitelino e a área pericárdica. Em ambos os casos, os gráficos indicaram um aumento significativo para classe “alterado”, o que pode ser atribuído à presença de embriões com edema nesses órgãos (Figura 11C-D). Além disso, observa-se nos gráficos uma grande dispersão de valores para ambos os parâmetros, demonstrando a alta variabilidade morfológica encontrada nesta classe. Em alguns casos, embriões pertencentes ao grupo “alterado” apresentaram características morfológicas semelhantes ao grupo normal, o que pode indicar variações fenotípicas menos pronunciadas.

Embora tenha sido possível distinguir os fenótipos a partir de alguns parâmetros extraídos pelo CellProfiler, eles podem não ser suficiente para a identificação de alterações morfológicas mais complexas. Diante disso, optou-se por empregar uma abordagem de

análise multiparamétrica baseada em inteligência artificial, visando uma melhor distinção entre os diferentes grupos fenotípicos. Para isso, o modelo de aprendizado de máquina XGBoost foi selecionado em função de alguns fatores, dentre eles a alta performance em tarefas de classificação, pela capacidade de resolver internamente situações onde há ausência de dados, por manejar grandes volumes de dados e apresentar robustez frente ao *overfitting*, em virtude dos mecanismos de regularização e de flexibilidade proporcionada por seus hiperparâmetros ajustáveis (Wiens *et al.*, 2025). Isto foi particularmente importante no nosso estudo já que nem todas as imagens possuíam valores para alguns parâmetros extraídos pelo CellProfiler, como no caso da área pericárdica e do saco vitelino em embriões não eclodidos. Ao final do treinamento o modelo apresentou um bom desempenho tanto no conjunto de treinamento (MCC = 0.91) quanto de teste (MCC = 0.84), uma diferença cerca de 7% entre os dois conjuntos (Figura 12A). Além disso, outras métricas comumente usadas para avaliar a performance de classificadores também apresentaram valores por volta de 0,9 no conjunto teste, corroborando a boa performance do modelo.

Ao analisar a matriz de confusão do modelo de melhor desempenho (i.e. maior valor de MCC) é possível notar taxas de acerto de 100% para embriões não eclodidos, 97% para embriões normais e 78% para o grupo alterado (Figura 12B). Nas outras versões do modelo, apenas a taxa de acerto da classe “alterado” se mostrou consideravelmente mais baixa, alcançando 65% (Figura A4). Esses resultados evidenciam a eficácia do modelo na identificação de fenótipos distintos, especialmente na detecção de embriões não eclodidos e normais. A menor taxa de acerto no grupo “alterado” deve-se, principalmente, a classificação errada de embriões de fenótipo “alterado” como “normal”. Isto pode ter ocorrido devido à sutileza de algumas alterações presentes no fenótipo “alterado” que, embora sejam visíveis, não foram detectados pelo modelo (Figura 12C). É provável a incorporação de mais imagens com fenótipos discretos no treinamento do modelo possa ajudar a obter um classificador com melhor performance.

O próximo passo deste trabalho é realizar novos ensaios de toxicidade com o modelo *zebrafish*, a fim de expandir o banco de dados e aprimorar o desempenho dos modelos de inteligência artificial desenvolvidos. A maior variabilidade fenotípica obtida por meio dessas novas imagens também permitirá o refinamento do classificador, com a inclusão de novas classes de fenótipos associados a alterações morfológicas específicas, como edema do saco vitelino, edema da área pericárdica, lordose e ausência de somitos. Por fim, realizaremos uma análise comparativa entre o nosso protocolo automatizado e o método da microscopia clássica,

com intuito de avaliar a sua real aplicabilidade na detecção de compostos com atividade tóxica.

## 5 CONCLUSÃO

Neste trabalho iniciamos o processo de padronização de um novo ensaio de HCA para avaliar a toxicidade de compostos químicos sobre embriões de zebrafish. Apesar dos experimentos não terem sido validados, as imagens geradas foram utilizadas para criar um modelo de segmentação e treinar o modelo de classificadores baseados em parâmetros extraídos do CellProfiler. A segmentação automática realizada por meio da rede neural convolucional (Resnet 34) apresentou desempenho variável de acordo com a região analisada. Enquanto a segmentação do corpo inteiro mostrou-se mais precisa (Dice = 0,95), regiões menores e menos contrastadas, como saco vitelino (Dice = 0,80) e área pericárdica (Dice = 0,25), evidenciaram limitações, possivelmente por conta da complexidade morfológica dessas estruturas, assim como ao número reduzido de imagens, resultando na ocorrência de possível *overfitting*. Também foi possível relacionar alguns fenótipos dos embriões com parâmetros do CellProfiler, em especial o “form factor”, calculado a partir da segmentação do corpo inteiro, que apresentam valores elevados (próximo de 1) apenas nos embriões não eclodidos. Já o classificador fenotípico XGBoost apresentou uma boa performance (MCC = 0,84), sendo muito eficaz (>90% de acerto) na detecção de embriões com fenótipo “normal” e “não eclodido”, mas com menor capacidade de identificar indivíduos “alterados” (65-78% de acerto). Embora os modelos baseados em aprendizado de máquina ainda não estejam plenamente aptos para aplicações automáticas nos ensaios atuais, eles ainda podem ser otimizados através da captura de mais imagens e ajustes de outros parâmetros relevantes para o treinamento. Por fim, é imprescindível a realização de novos experimentos para avaliar a capacidade do nosso ensaio na detecção de compostos com atividade tóxica conhecida, permitindo uma comparação direta com os resultados obtidos por métodos tradicionais de microscopia.

## REFERÊNCIAS

- AKERBERG, A. A. *et al.* Deep learning enables automated volumetric assessments of cardiac function in zebrafish. **Disease Models & Mechanisms**, v. 12, n. 10, out. 2019.
- AZEVEDO, L. D. *et al.* Spirooxadiazoline-oxindoles derived from imatinib show antimyeloproliferative potential in K562 cells. **Archiv der Pharmazie**, v. 357, n. 8, ago. 2024.
- BAUER, B.; MALLY, A.; LIEDTKE, D.. Zebrafish Embryos and Larvae as Alternative Animal Models for Toxicity Testing. **International journal of molecular sciences**, v. 22, n. 24, p. 13417, dez. 2021.
- BECKMAN, S. Automated Image Screening of Zebrafish Embryos Exposed to Developmental Toxins. **Developmental Biology**, 2018.
- BRASIL. Ministério da Ciência, Tecnologia, Inovações e Comunicações. Resolução CONCEA nº 37, de 15 de fevereiro de 2018. Disponível em: [https://antigo.mctic.gov.br/mctic/opencms/legislacao/outros\\_atos/resolucoes/Resolucao\\_CONCEA\\_n\\_37\\_de\\_15022018.html](https://antigo.mctic.gov.br/mctic/opencms/legislacao/outros_atos/resolucoes/Resolucao_CONCEA_n_37_de_15022018.html). Acesso em: 11 feb. 2025
- BRAY, M. A. *et al.* Advanced Assay Development Guidelines for Image-Based High Content Screening and Analysis. **Assay Guidance Manual**. jul. 2017
- BUCHSER, W. *et al.* Assay Development Guidelines for Image-Based High Content Screening, High Content Analysis and High Content Imaging. **Assay Guidance Manual**. oct. 2012
- CANNY, S.G.T. *et al.* Go Fish! Developmental Stages in Zebrafish. **Genesis**, v 47, mar. 2009
- CASSAR, Steven *et al.* Use of Zebrafish in Drug Discovery Toxicology. **Chemical Research in Toxicology**, v. 33, n. 1, p. 95–118, 21 jan. 2020.
- CASSAR, S.; HUANG, X.; COLE, T.. A high-throughput method for predicting drug effects on gut transit time using larval zebrafish. **Journal of Pharmacological and Toxicological Methods**, v. 76, p. 72–75, nov. 2015.
- CHAPELA, D. *et al.* A zebrafish drug screening platform boosts the discovery of novel therapeutics for spinal cord injury in mammals. **Scientific Reports**, v. 9, n. 1, p. 10475, 19 jul. 2019.

CHEN, S. *et al.* A high-throughput screening assay for identification of chemicals with liver tumor promoting potential using a transgenic zebrafish line. **Chemosphere**, v. 297, p. 134169, jun. 2022.

CHICCO, D. *et al.* The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. **BMC Genomics**, v. 21, n. 6, jan. 2020

CULLY, M. Zebrafish earn their drug discovery stripes. **Nature Reviews Drug Discovery**, v. 18, n. 11, p. 811–813, 3 out. 2019.

DEAL, S. *et al.* Development of a quantitative morphological assessment of toxicant-treated zebrafish larvae using brightfield imaging and high-content analysis. **Journal of Applied Toxicology**, v. 36, n. 9, p. 1214–1222, set. 2016.

DUAN, J. *et al.* Low-dose exposure of silica nanoparticles induces cardiac dysfunction via neutrophil-mediated inflammation and cardiac contraction in zebrafish embryos. **Nanotoxicology**, v. 10, n. 5, p. 575–585, 27 may. 2016.

GARCIA, G. R.; NOYES, P. D.; TANGUAY, R. L. Advancements in zebrafish applications for 21st century toxicology. **Pharmacology & Therapeutics**, v. 161, p. 11–21, may. 2016.

GOMES, B. F. *et al.* Discovery of new *Schistosoma mansoni* aspartyl protease inhibitors by structure-based virtual screening. **Memórias do Instituto Oswaldo Cruz**, v. 118, sep. 2023.

HELLFELD, R. V. *et al.* Adverse effects in the fish embryo acute toxicity (FET) test: a catalogue of unspecific morphological changes versus more specific effects in zebrafish (*Danio rerio*) embryos. **Environmental Sciences Europe**, v. 32, n. 122, sep. 2020.

HORIE, Y. *et al.* Lethal and sublethal effects of aniline and chlorinated anilines on zebrafish embryos and larvae. **Journal of Applied Toxicology**, v. 37, n. 7, p. 836–841, jul. 2017.

HOYBERGHS, J. *et al.* DMSO Concentrations up to 1% are Safe to be Used in the Zebrafish Embryo Developmental Toxicity Assay. **Frontiers in Toxicology**, v. 3, dez. 2021.

JEANRAY, NATHALIE *et al.* Phenotype Classification of Zebrafish Embryos by Supervised Learning. **PLOS ONE**, v. 10, n. 1, p. e0116989, 9 jan. 2015.

KHALILI, A.; REZAI, P. Microfluidic devices for embryonic and larval zebrafish studies. **Briefings in Functional Genomics**, v. 18, n. 6, p. 419–432, 19 nov. 2019.

- LAMPRECHT, M. R.; SABATINI, D. M.; CARPENTER, A. E. CellProfiler™: Free, Versatile Software for Automated Biological Image Analysis. **BioTechniques**, v. 42, n. 1, p. 71–75, jan. 2007.
- LEYDEN, C. *et al.* Efficacy of Tricaine (MS-222) and Hypothermia as Anesthetic Agents for Blocking Sensorimotor Responses in Larval Zebrafish. **Frontiers in Veterinary Science**, v. 9, mar. 2022.
- LIU, H. *et al.* A High-Content Larval Zebrafish Brain Imaging Method for Small Molecule Drug Discovery. **PLOS ONE**, v. 11, n. 10, p. e0164645, out. 2016.
- LOBATO, H. *et al.* Prediction of long-term creep modulus of thermoplastics using brief tests and interpretable machine learning. **International Journal of Solids and Structures**, v. 304, nov. 2024.
- MISCEVIC, F.; ROTSTEIN, O.; WEN, X.. Advances in Zebrafish High Content and High Throughput Technologies. **Combinatorial Chemistry & High Throughput Screening**, v. 15, n. 7, p. 515–521, 1 jul. 2012.
- OECD. **Test No. 236: Fish Embryo Acute Toxicity (FET) Test**. [*S.l.*]: OECD, 2013.
- OTTERSTROM, J. J. *et al.* Technologies bringing young Zebrafish from a niche field to the limelight. **SLAS Technology**, v. 27, n. 2, p. 109–120, abr. 2022.
- PADILLA, S.; GLABERMAN, S.. The zebrafish (*Danio rerio*) model in toxicity testing. *In: An Introduction to Interdisciplinary Toxicology*. [*S.l.*]: Elsevier, p. 525–532, jan. 2020.
- PATTON, E. E.; ZON, L. I.; LANGENAU, DAVID M. Zebrafish disease models in drug discovery: from preclinical modelling to clinical trials. **Nature Reviews Drug Discovery**, v. 20, n. 8, p. 611–628, 11 ago. 2021.
- REID, R. M.; D'AQUILA, A. L.; BIGA, P. R. The validation of a sensitive, non-toxic in vivo metabolic assay applicable across zebrafish life stages. **Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology**, v. 208, p. 29–37, jun. 2018.
- RENQUIST, B. J. *et al.* Development of an Assay for High-Throughput Energy Expenditure Monitoring in the Zebrafish. **Zebrafish**, v. 10, n. 3, p. 343–352, set. 2013.

SARAMAGO, L. C. *et al.* AI-Driven Discovery of SARS-CoV-2 Main Protease Fragment-like Inhibitors with Antiviral Activity *In Vitro*. **Journal of Chemical Information and Modeling**, v. 63, n. 9, p. 2866–2880, maio 2023.

SENG, W. L.; ZHANG, D.; MCGRATH, P.. Microplate-Based Whole Zebrafish Caspase 3/7 Assay for Screening Small Molecule Compounds. *In: [S.l.: S.n.]*. p. 193–209, apr. 2016.

SENGER, M. R. *et al.* Kinetics and molecular modeling studies on the inhibition mechanism of GH13  $\alpha$ -glycosidases by small molecule ligands. **International Journal of Biological Macromolecules**, v. 269, p. 132036, jun. 2024.

SHANG, S.; LIN, S.; CONG, F. Zebrafish Larvae Phenotype Classification from Bright-field Microscopic Images Using a Two-Tier Deep-Learning Pipeline. **Applied Sciences**, v. 10, n. 4, p. 1247, fev. 2020.

TAHA, A. A.; HANBURY, A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. **BMC Medical Imaging**, v. 15, n. 1, p. 29, dez. 2015.

TEIXIDÓ, E. *et al.* Automated Morphological Feature Assessment for Zebrafish Embryo Developmental Toxicity Screens. *Toxicological Sciences*, v. 167, p. 438-449, feb. 2019.

WESTERFIELD, M. The zebrafish book. A guide for the laboratory use of zebrafish (*Danio rerio*). 5 ed., Univ. of Oregon Press, Eugene, 2007.

WESTHOFF, J. H. *et al.* In vivo High-Content Screening in Zebrafish for Developmental Nephrotoxicity of Approved Drugs. **Frontiers in Cell and Developmental Biology**, v. 8, jul. 2020.

WIENS, M. *et al.* A Tutorial and Use Case Example of the eXtreme Gradient Boosting (XGBoost) Artificial Intelligence Algorithm for Drug Development Applications. **Clinical and Translational Science**, v. 18, mar. 2025.

YANG, X. *et al.* High-throughput Screening in Larval Zebrafish Identifies Novel Potent Sedative-hypnotics. **Anesthesiology**, v. 129, n. 3, p. 459–476, set. 2018.

YOZZO, K. *et al.* High-Content Screening Assay for Identification of Chemicals Impacting Cardiovascular Function in Zebrafish Embryos. **Environmental Science & Technology**, v. 47, n. 19, p. 11302–11310, out. 2013.

ZHANG, B. *et al.* Quantitative Phenotyping-Based In Vivo Chemical Screening in a Zebrafish Model of Leukemia Stem Cell Xenotransplantation. **PLoS ONE**, v. 9, n. 1, p. e85439, jan. 2014.

ZHANG, G. *et al.* An improved automated zebrafish larva high-throughput imaging system. **Computers in Biology and Medicine**, v. 136, p. 104702, set. 2021.

ZHU, X. *et al.* Fenobucarb-induced developmental neurotoxicity and mechanisms in zebrafish. **NeuroToxicology**, v. 79, p. 11–19, jul. 2020.

## ANEXOS

Hiperparâmetro	Intervalo proposto	Melhores parâmetros (corpo inteiro)	Melhores parâmetros (saco vitelino)	Melhores parâmetros (área pericárdica)
epochs	[5, 100]	35	93	11
freeze_epochs	[1, 5]	1	3	5
base_lr	[0.0001, 0.01]	0.002	0.001	0.0002
wd (weight decay)	[1e-6, 1e-5, 1e-4, 1e-3, 1e-2]	None	1e-05	0.01

Tabela A1: Hiperparâmetros do modelo Resnet 34 otimizados pelo optuna. Na tabela são mostrados os intervalos propostos pelo usuário para cada hiperparâmetro e os melhores valores encontrados pelo optuna para a versão do modelo de maior desempenho.

Hiperparâmetro	Intervalo proposto	Melhores parâmetros
max_depth	[2, 8]	2
gamma	[0.5, 5.0]	5.0
reg_alpha	[4, 10]	0.2
reg_lambda	[1e -4, 10]	0.0008
subsample	[0.5, 1.0]	0.8
colsample_bytree	[0.5, 1.0]	0.55
min_child_weight	[3, 10]	9
n_estimators	[300, 1500]	1350
eta	[1e -4, 0.05]	0.001

Tabela A2: Hiperparâmetros do modelo de classificação XGBoost otimizados pelo optuna. A tabela mostra intervalos fornecidos pelo usuário para cada hiperparâmetro e os melhores valores encontrados pelo optuna para a versão do modelo de maior desempenho.

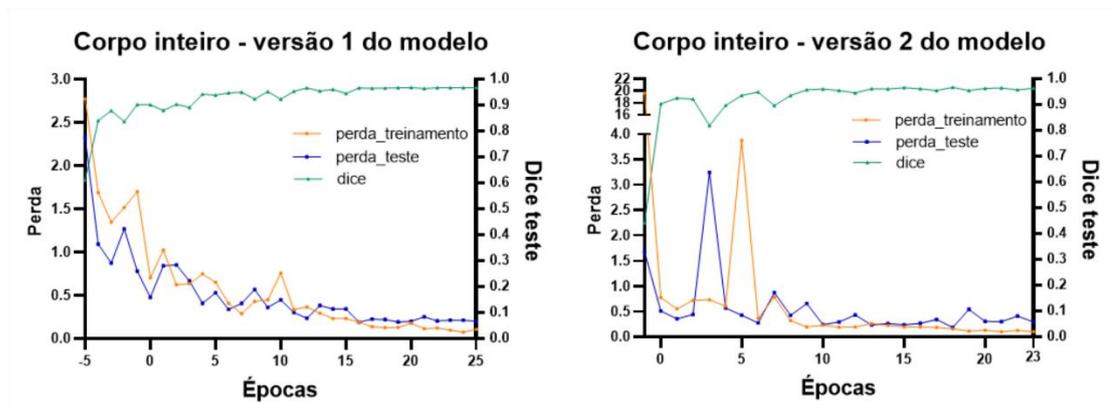


Figura A1: Curvas de perda das demais versões do modelo de segmentação do corpo inteiro.

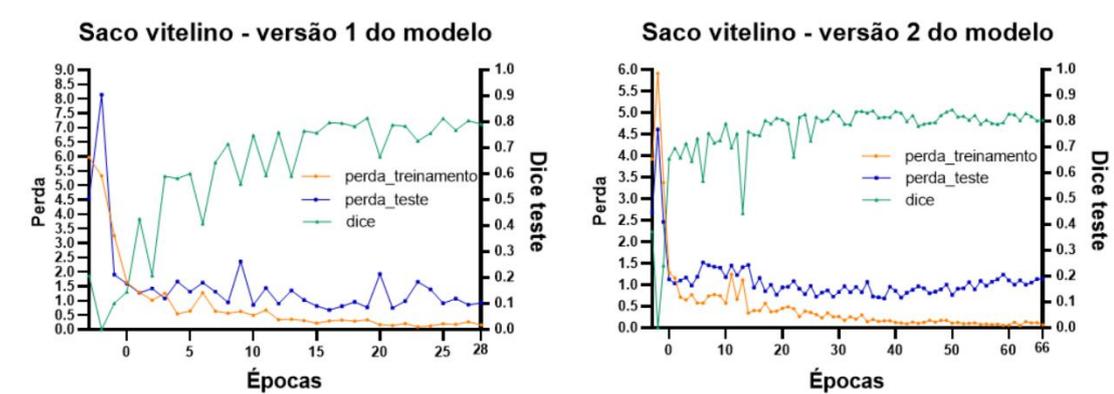


Figura A2: Curvas de perda das demais versões do modelo de segmentação do saco vitelino.

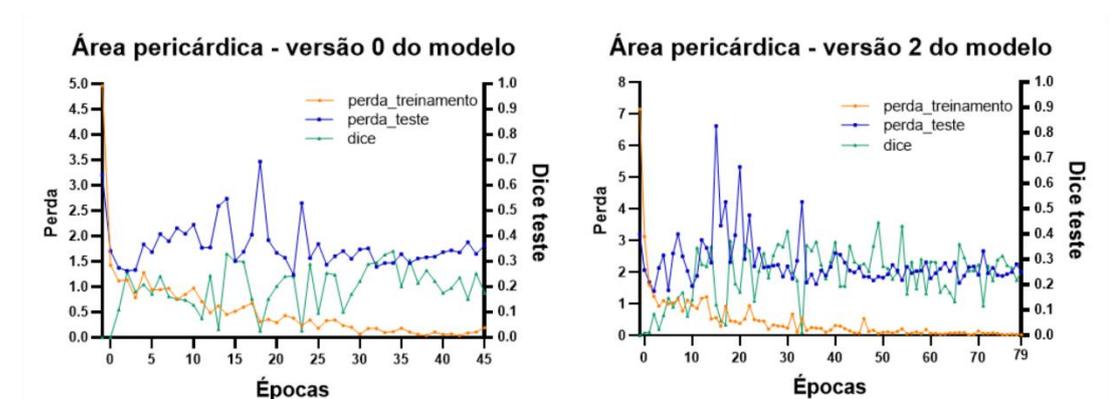


Figura A3: Curvas de perda e valor de Dice das demais versões do modelo de segmentação da área pericárdica.

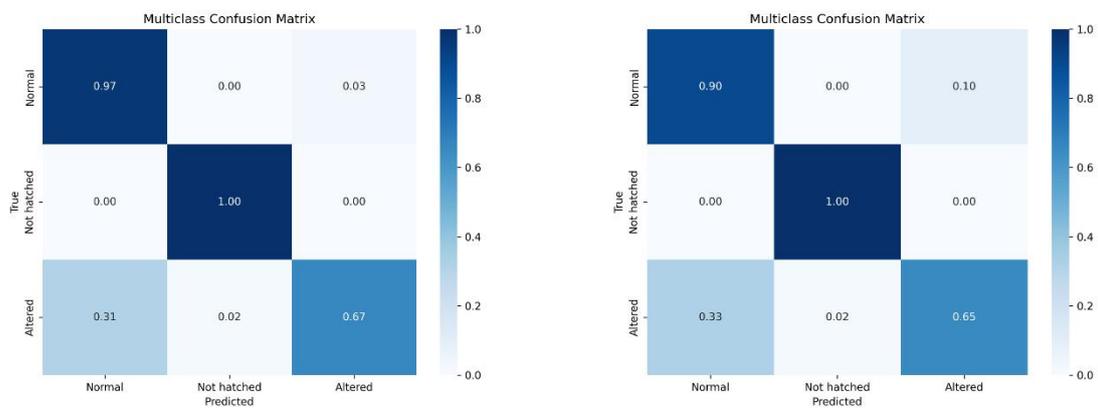


Figura A4: Matrizes de confusão das demais versões do modelo XGBoost.