

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

GUILHERME PORTO DO NASCIMENTO

**PROCESSAMENTO DE IMAGENS PARA A EXTRAÇÃO DE PARÂMETROS DE
ESCOAMENTOS BIFÁSICOS GÁS-LÍQUIDO UTILIZANDO REDES NEURAIS
CONVOLUCIONAIS**

CURITIBA

2026

GUILHERME PORTO DO NASCIMENTO

**PROCESSAMENTO DE IMAGENS PARA A EXTRAÇÃO DE PARÂMETROS DE
ESCOAMENTOS BIFÁSICOS GÁS-LÍQUIDO UTILIZANDO REDES NEURAIAS
CONVOLUCIONAIS**

**Image Processing of Gas–Liquid Two-Phase Flow for Parameter Extraction
Using Convolutional Neural Networks**

Dissertação apresentada como requisito para obtenção do título de Mestre em Ciências, do Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, da Universidade Tecnológica Federal do Paraná (UTFPR). Área de Concentração: Engenharia de Computação.

Orientador: Prof. Dr. Marco José da Silva
Coorientador: Prof. Dr. Eduardo Nunes dos Santos

CURITIBA

2026



Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Curitiba



GUILHERME PORTO DO NASCIMENTO

**PROCESSAMENTO DE IMAGENS PARA A EXTRAÇÃO DE PARÂMETROS DE ESCOAMENTOS BIFÁSICOS
GÁS-LÍQUIDO UTILIZANDO REDES NEURAIAS CONVOLUCIONAIS**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Ciências da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Engenharia De Automação E Sistemas.

Data de aprovação: 19 de Fevereiro de 2026

Dr. Aluisio Do Nascimento Wrasse, Doutorado - Instituto Federal de Educação, Ciência e Tecnologia Sul-Rio-Grandense (Ifsul)

Dr. Daniel Rodrigues Pipa, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Eduardo Nunes Dos Santos, Doutorado - Universidade Tecnológica Federal do Paraná

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 19/02/2026.

AGRADECIMENTOS

Agradeço ao meu orientador e ao meu coorientador, pela orientação, paciência e apoio ao longo da elaboração deste trabalho, assim como à Coordenação do Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, pela cooperação, compreensão e prestatividade ao longo dessa trajetória acadêmica.

Aos meus colegas de laboratório, agradeço pela convivência, troca de conhecimento e pelo apoio prestado durante o desenvolvimento do trabalho. Além da Universidade Tecnológica Federal do Paraná (UTFPR) e o Núcleo de Escoamento Multifásico (NUEM), por proporcionar as oportunidades necessárias para que fosse possível dar início a minha carreira acadêmica.

RESUMO

Escoamentos multifásicos são amplamente encontrados em diversos setores industriais, com destaque para as indústrias nuclear, química e de óleo e gás. Esses escoamentos são compostos por duas ou mais fases, que podem assumir diferentes configurações geométricas em função das propriedades físicas das fases e das condições operacionais. Na produção de petróleo e gás, os escoamentos são predominantemente compostos por água, óleo e gás natural. No cenário brasileiro, o pré-sal é responsável por grande parte da produção brasileira de petróleo e gás, correspondendo a aproximadamente 80% do total, conforme dados divulgados pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP). A exploração desses reservatórios ocorre majoritariamente em águas profundas, caracterizada por elevadas pressões e altos teores de CO₂, características que impactam diretamente o comportamento do escoamento e exigem novas estratégias de projeto, operação e monitoramento dos sistemas de produção. A obtenção de informações experimentais sobre esse fenômeno é fundamental para o estabelecimento de correlações de engenharia e para o desenvolvimento de modelos que descrevem o seu comportamento físico. Este trabalho apresenta uma metodologia baseada em técnicas de processamento de imagens para a extração de parâmetros característicos do escoamento bifásico gás-líquido em regime de golfadas. Os dados experimentais são obtidos por meio de uma bancada com temperatura e pressão controladas, utilizando fluidos modelo (óleo mineral e hexafluoreto de enxofre) para emular condições de campo. As imagens adquiridas em altas taxas de aquisição são segmentadas por meio de uma rede neural convolucional do tipo U-Net, possibilitando a distinção entre as regiões de líquido e de gás, com coeficiente *Dice* médio superior a 0,95. As bolhas alongadas são identificadas e reconstruídas através da agregação de frames consecutivos, com base na velocidade do nariz da bolha. Os parâmetros que caracterizam o escoamento estudado são estimados e comparados com modelos clássicos da literatura, medições manuais e técnicas baseadas em sensores capacitivos, apresentando desvios da ordem de $\pm 10\%$ em relação às referências adotadas. A metodologia proposta demonstra ser capaz de estimar, de forma automática e robusta, os principais parâmetros característicos do escoamento bifásico em regime de golfadas.

Palavras-chave: escoamento bifásico gás-líquido; regime de golfadas; processamento de imagens; redes neurais convolucionais.

ABSTRACT

Multiphase flows are widely encountered in several industrial sectors, particularly in the nuclear, chemical, and oil and gas industries. These flows are composed of two or more phases, which may assume different geometric configurations depending on the physical properties of the phases and the operating conditions. In oil and gas production, the flows are predominantly composed of water, oil, and natural gas. In the Brazilian context, the pre-salt reservoirs account for a large share of the country's oil and gas production, corresponding to approximately 80% of the total, according to data released by the National Agency of Petroleum, Natural Gas and Biofuels (ANP). The exploitation of these reservoirs occurs mainly in deepwater environments, characterized by high pressures and high CO₂ contents, features that directly affect flow behavior and require new strategies for the design, operation, and monitoring of production systems. The acquisition of experimental information on this phenomenon is essential for establishing engineering correlations and for developing models that describe its physical behavior. This work presents a methodology based on image processing techniques for the extraction of characteristic parameters of gas–liquid two-phase slug flow. Experimental data are obtained using a test facility with controlled temperature and pressure, employing model fluids (mineral oil and sulfur hexafluoride) to emulate field conditions. Images acquired at high acquisition rates are segmented using a U-Net convolutional neural network, enabling the distinction between liquid and gas regions, with an average Dice coefficient higher than 0.95. Elongated bubbles are identified and reconstructed through the aggregation of consecutive frames, based on the bubble nose velocity. The parameters that characterize the studied flow are estimated and compared with classical models from the literature, manual measurements, and techniques based on capacitive sensors, showing deviations on the order of $\pm 10\%$ relative to the adopted references. The proposed methodology demonstrates the capability to automatically and robustly estimate the main characteristic parameters of gas–liquid two-phase slug flow.

Keywords: gas–liquid two-phase flow; slug flow; image processing; convolutional neural networks.

LISTA DE ILUSTRAÇÕES

Figura 1 – Padrões de escoamento em tubulações horizontais e quase horizontais	20
Figura 2 – Padrões de escoamento em tubulações verticais e fortemente inclinadas.....	22
Figura 3 – Representação da célula unitária: (a) Célula unitária; (b) Seção transversal da bolha alongada; (c) Seção transversal do pistão de líquido	25
Figura 4 – Mapa de escoamento bifásico líquido-gás.....	28
Figura 5 – Imagem em escala de cinza de uma tubulação horizontal em um sistema de escoamento bifásico	30
Figura 6 – Exemplo de segmentação com limiar de intensidade.....	32
Figura 7 – Elementos estruturantes e conversão em arranjos retangulares	34
Figura 8 – Exemplos de operações morfológicas em máscara binária.....	36
Figura 9 – Exemplo de aplicação do método watershed	37
Figura 10 – Detecção de bordas em imagens de escoamento bifásico utilizando os operadores Sobel e Canny	39
Figura 11 - Segmentação da imagem utilizando o algoritmo K-means	41
Figura 12 – Exemplo de convolução	45
Figura 13 – Diagrama arquitetura U-Net.....	47
Figura 14 – Diagrama do circuito experimental.....	59
Figura 15 – Janela de visualização: (a) Esquemático; (b) Janela de visualização com iluminação ligada.....	60
Figura 16 - Exemplos de segmentação	61
Figura 17 – Diagrama da arquitetura final do modelo adotado	62
Figura 18 – Distribuição de velocidades superficiais	63
Figura 19 – (a) Frame segmentado com detecções e caixas delimitadoras; (b) Matriz representando o conjunto de trilhas associadas entre frames sucessivos	68
Figura 20 - Deslocamento do nariz da bolha em frames consecutivos.....	71
Figura 21 – Exemplo de imagens agregadas	73
Figura 22 – Representação da célula unitária e medidas geométricas	74
Figura 23 – Representação da célula unitária e medidas geométricas	75
Figura 24 – Velocidade de bolhas alongadas para bateria de 3 bar	77
Figura 25 – Frequência de bolhas alongadas para bateria de 3 bar	77
Figura 26 – Exemplos de imagens correspondentes aos pontos fora do intervalo de $\pm 10\%$ no gráfico de frequência (3 bar).....	78

Figura 27 – Teste com diferentes conjuntos de dados. (a) Dice médio para máscaras manuais e semiautomáticas em função do tamanho do conjunto de dados. (b) Dice por subconjunto para validação cruzada das máscaras semiautomáticas. (c) Dice por subconjunto para validação cruzada das máscaras anotadas manualmente	79
Figura 28 – Comparação entre a imagem original, a máscara de referência (<i>ground truth</i>) e a máscara segmentada pela U-Net.	80
Figura 29 – Modelo clássico de velocidade de bolha alongada.....	82
Figura 30 – Número de Strouhal em função da fração de velocidade superficial do líquido	83
Figura 31 – Estimativa de fração de vazio em função da fração de vazio estimada experimentalmente	84
Figura 32 – Comparação entre os parâmetros estimados por sensor capacitivo e por câmera	86
Figura 33 – Histograma da velocidade de bolha.....	100
Figura 34 – Histograma da fração de vazio.....	101
Figura 35 – Histograma do comprimento de célula unitária.....	102
Figura 36 – Histograma de comprimento de bolha.....	103
Figura 37 – Histograma de comprimento de pistão de líquido.....	104

LISTA DE TABELAS

Tabela 1 - Comparação entre estudos que aplicaram U-Net à segmentação em escoamentos bifásicos	55
Tabela 2 – Bateria experimental a 3 bar	56
Tabela 3 – Métricas de erro da estimativa de fração de vazio	85
Tabela 4 - RMSE e MAPE obtidos para os parâmetros estimados pela câmera em comparação com os sensores capacitivos	87
Tabela 5 – Baterias experimentais a 4 bar	98

LISTA DE ABREVIATURAS E SIGLAS

UTFPR	Universidade Tecnológica Federal do Paraná
NUEM	Núcleo de Escoamento Multifásico
CO ₂	Dióxido de carbono
SF ₆	Hexafluoreto de enxofre
SG	<i>Shadowgraphy</i>
PIV	<i>Particle Image Velocimetry</i>
FI	<i>Fluorescence Imaging</i>
SRG	<i>Seeded Region Growing</i>
IIC	<i>Invariant Information Clustering</i>
SAM	<i>Segment Anything Model</i>
ANN	<i>Artificial Neural Network</i>
MLP	<i>Multilayer Perceptron</i>
CNN	<i>Convolutional Neural Network</i>
IoU	<i>Intersection over Union</i>
VP	Verdadeiro positivo
FP	Falso positivo
VN	Verdadeiro negativo
FN	Falso negativo
DBT	<i>Detection-Based Tracking</i>
MOT	<i>Multiple Object Tracking</i>
RMSE	<i>Root Mean Square Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>

LISTA DE SÍMBOLOS

W	Vazão mássica
Q	Vazão volumétrica
J	Velocidade superficial
A	Área da seção transversal
J _G	Velocidade superficial de gás
J _L	Velocidade superficial de líquido
Q _G	Vazão volumétrica de líquido
Q _L	Vazão volumétrica de gás
H _L	Fração de líquido
α	Fração de gás
U _B	Velocidade de bolha alongada
L _S	Comprimento de pistão de líquido
L _B	Comprimento de bolha alongada
h _L	Altura de líquido
f	Frequência
C ₀	Parâmetro de distribuição
u _∞	Velocidade de deriva
g	Aceleração da gravidade
D	Diâmetro da tubulação
ρ_G	Densidade do gás
ρ_L	Densidade do líquido
C _∞	Constante de deslizamento
S _t	Número de Strouhal
H _{LS}	Fração de líquido na região do pistão
V _{LS}	Volume de líquido na região do pistão
V _S	Volume total da região do pistão
H _{LB}	Fração de líquido na região da bolha
V _{LB}	Volume de líquido na região da bolha
V _B	Volume total da região da bolha

SUMÁRIO

1	<i>Introdução</i>	13
1.1	Motivação	13
1.2	Objetivo	15
1.3	Estrutura	16
2	<i>Revisão da literatura</i>	18
2.1	Escoamento bifásico gás-líquido	18
2.1.1	Padrões de escoamento.....	18
2.1.2	Parâmetros de escoamento	22
2.1.3	Mapa de padrões de escoamento	27
2.2	Processamento de imagens	28
2.2.1	Conceitos básicos	28
2.2.2	Segmentação de imagens.....	30
2.3	Rastreamento de objetos	49
2.4	Trabalhos relacionados	53
3	<i>Materiais e métodos</i>	56
3.1	Circuito experimental	58
3.2	Segmentação de imagens	61
3.3	Rastreamento de objetos	66
3.4	Parâmetros de escoamento	70
4	<i>Resultados</i>	76
4.1	Segmentação de imagens	78
4.2	Parâmetros de escoamento	81
5	<i>Conclusão</i>	88
	REFERÊNCIAS	90
	APÊNDICE A - Tabelas complementares das baterias experimentais	
	97	
	APÊNDICE B - Histogramas em matriz dos parâmetros do	
	escoamento (4 bar)	99

1 INTRODUÇÃO

1.1 Motivação

Os escoamentos multifásicos estão presentes em diversos setores industriais, incluindo as indústrias nuclear, química e de petróleo e gás. Esses escoamentos são compostos por uma mistura de duas ou mais fases, cujo comportamento é influenciado por uma variedade de fatores. Dentre esses fatores, destacam-se as propriedades dos fluidos, as interações entre as fases, as condições operacionais e a geometria dos sistemas de escoamento (Shoham, 2006). No cenário específico da produção de petróleo e gás, esses escoamentos são compostos por óleo, água, gás natural e sedimentos. No contexto brasileiro, de acordo com dados da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), o pré-sal é responsável por mais de 80% da produção nacional de petróleo e gás. A exploração dessas reservas envolve desafios significativos, tais como altas pressões, baixas temperatura, elevados teores de CO₂ e uma menor razão de densidades entre a fase gasosa e líquida (Beltrão et al., 2009). Essas características podem alterar os mecanismos de formação, transporte e estabilidade das estruturas do escoamento quando comparada a sistemas que não apresentam tais particularidades. Dessa forma, torna-se evidente a necessidade de estudos experimentais que incorporem as características específicas dos escoamentos em águas profundas do pré-sal, uma vez que a maioria dos modelos, correlações de engenharia e mapas de padrões escoamento foram desenvolvidos a partir de estudos que não as consideram.

Embora o escoamento multifásico possa envolver diferentes combinações de fases, grande parte dos fenômenos hidrodinâmicos pode ser descrita adequadamente por meio de uma abordagem bifásica. Trabalhos recentes apresentam uma caracterização experimental para o estudo utilizando fases como hexafluoreto de Kadri Kadrienxofre (SF₆) e óleo mineral (Kadri; Mudde; Oliemans, 2010; Gmyterco, 2018; Quintino; Da Fonseca Junior; Rodriguez, 2023). A escolha se dá por apresentarem propriedades que emulam o comportamento dos fluidos nas condições do pré-sal, mantendo a razão das massas específicas mesmo com pressões de operação reduzidas. A motivação central desses trabalhos reside na busca de ampliar e validar modelos e simuladores de escoamento bifásico em regimes com elevada densidade de gás e baixa razão de densidade (gás com densidade relativamente próxima a do líquido).

A análise dos parâmetros de escoamento bifásico gás-líquido é uma forma eficiente de caracterizar os padrões de escoamento (Shoham, 2006), validar modelos propostos e fornecer dados para o desenvolvimento de correlações de engenharia. Sendo assim, diversas tecnologias de medição são utilizadas com este propósito, incluindo sensores de impedância elétrica, ópticos, ultrassônicos, raio-X, raio- γ e impedância elétrica (Martins, 2022), com foco em medição de parâmetros como velocidade, comprimento e frequência de bolhas alongadas. Diversos estudos estabeleceram os parâmetros característicos a serem calculados em escoamento bifásicos gás-líquido, bem como simplificações e procedimentos necessários para sua determinação. Entre eles, destacando-se os estudos de Bendiksen (1984), Taitel e Barnea (1990) e Kadri, Mudde e Oliemans (2010).

Ademais, outra abordagem amplamente utilizada em experimentos envolvendo dinâmica de fluidos é a captura imagens de alta velocidade. Esse método permite uma visualização detalhada através da aquisição de quadros de imagens com uma alta resolução espacial e temporal. As câmeras de alta velocidade utilizadas para aquisição podem variar de dezenas até milhões de quadros por segundos, dependendo da resolução da imagem adquirida, da iluminação do ambiente e da capacidade do sensor e armazenamento da câmera. Em contraste, as câmeras possuem característica de serem não intrusivas, demandando apenas uma janela de visualização e uma fonte de iluminação adequada para capturar os fenômenos sem interferir no sistema. Existem diversas técnicas para aquisição de imagens de alta velocidade, entre elas *Shadowgraphy* (SG), *Particle Image Velocimetry* (PIV) e *Fluorescence Imaging* (FI). Essas técnicas diferem em aspectos como o tipo de iluminação, a necessidade de filtros ópticos e o uso de corantes ou partículas de rastreamento (Versluis, 2012).

Após as imagens serem adquiridas, um dos principais desafios consiste na identificação precisa das fases, uma vez que é necessário superar empecilhos como ruídos, distorções ópticas, variações de iluminação e sobreposições. Alguns exemplos de trabalhos relacionados a escoamento bifásico envolvem técnicas de segmentação baseadas em morfologia matemática (Mayor; Pinto; Campos, 2007), no método *level set* (Coelho, 2011), no algoritmo Watershed (Do Amaral et al., 2013a) e em limiares de intensidade (Kong et al., 2019). Mais recentemente, técnicas que utilizam inteligência artificial, como redes neurais utilizando arquiteturas como U-Net (Ronneberger; Fischer; Brox, 2015), DeepLab (Chen et al., 2016) e Mask R-CNN (He

et al., 2018) e SAM 2 (Ravi et al., 2024), são capazes de aprender a segmentar imagens que possuam características complexas e não lineares, sendo mais sensíveis a características contextuais, como intensidade, formato, posição e textura dos objetos. Outra característica importante é a capacidade de automatização e generalização de processos de segmentação, eliminando a necessidade de ajustes manuais, como a escolha de valores limiares. Todos esses pontos positivos são acompanhados de alguns desafios, como a necessidade de um grande conjunto de dados para treinamento e a necessidade de requisitos computacionais elevados, que podem ser maiores ou menores dependendo da arquitetura escolhida.

Este trabalho propõe uma metodologia de processamento de imagens para extração de parâmetros críticos de escoamento em regime de golfadas de forma automatizada, baseado em técnicas de imageamento de alta velocidade e aprendizado profundo. A abordagem combina aquisição de imagens em diferentes condições de escoamento, segmentação de imagens por meio de redes neurais convolucionais e rastreamento das estruturas gasosas, permitindo a reconstrução de suas trajetórias e a visualização de estruturas com dimensões superiores a região de visualização. A metodologia foi avaliada a partir da comparação dos resultados obtidos com modelos clássicos de parâmetros da literatura e com técnicas alternativas para a extração de parâmetros.

1.2 Objetivo

Este trabalho tem como objetivo desenvolver um método automatizado baseado em processamento de imagens para extração e análise de parâmetros associados ao escoamento bifásico líquido-gás. Os objetivos específicos são:

- Realizar um estudo das técnicas de processamento de imagens aplicadas a escoamento bifásico gás-líquido em regime de golfadas.
- Consolidar um conjunto de dados experimentais compostos por imagens do escoamento;
- Desenvolver uma metodologia fundamentada em técnicas de imageamento de alta velocidade, aprendizado profundo e rastreamento das estruturas gasosas para extração de parâmetros de escoamento.

- Treinar, validar e aplicar um modelo de segmentação generalista baseado em aprendizado profundo, robusto frente a diferentes condições adversas, incluindo variações de iluminação, ruídos e diversidade geométrica das fases;
- Investigar o tamanho mínimo necessário conjunto de dados para o treinamento do modelo de segmentação, analisando o impacto dessa variável no seu desempenho;
- Estimar parâmetros críticos do escoamento bifásico gás-líquido em regime de golfadas, de interesse para o desenvolvimento de modelos mecanicistas.
- Validar o método proposto através da comparação de seus resultados com aqueles obtidos a partir de técnicas alternativas de caracterização dos parâmetros de escoamento e de modelos clássicos da literatura.

1.3 Estrutura

O texto está organizado em 5 capítulos de modo a apresentar um panorama geral do trabalho, uma fundação teórica básica necessária para seu entendimento, a metodologia utilizada e os resultados da pesquisa.

- Capítulo 1 – Introdução: Contextualiza o tema geral definindo a problemática central, justifica a relevância do trabalho e estabelece os objetivos a serem alcançados;
- Capítulo 2 – Revisão Bibliográfica: Explora a temática de escoamento bifásico gás-líquido, assim como as técnicas de processamento de imagem aplicadas a este fenômeno voltada a exploração de óleo e gás;
- Capítulo 3 – Materiais e Métodos: Descreve o desenvolvimento do sistema de processamento de imagens para extração de parâmetros, abrangendo o circuito experimental utilizado e equipamentos para aquisição de imagens de alta velocidade, a elaboração de um conjunto de dados para treinamento de um modelo de segmentação e o desenvolvimento de algoritmos para estimação de parâmetros a partir de imagens segmentadas;

- Capítulo 4 – Resultados: Apresenta os resultados de validação da técnica proposta, incluindo a avaliação das segmentações e da estimação dos parâmetros que caracterizam o escoamento;
- Capítulo 5 – Conclusão: Resume as contribuições do desenvolvimento da pesquisa. Discute também as vantagens e limitações da metodologia adotada, além de propor direções para pesquisas futuras que podem expandir ou aprofundar esta investigação.

2 REVISÃO DA LITERATURA

Este capítulo apresenta uma revisão bibliográfica estruturada em quatro seções. A primeira tem como foco revisar os fundamentos teóricos e práticos do escoamento bifásico gás-líquido em regime de golfadas. A segunda é dedicada a técnicas de processamento de imagem aplicadas ao estudo da dinâmica dos fluidos. A terceira seção trata dos métodos de rastreamento de objetos, com ênfase em aplicações baseadas em visão computacional. Por fim, a quarta seção apresenta uma revisão de trabalhos correlatos que integram o processamento de imagens à análise de escoamentos multifásicos em regime de golfadas. Em conjunto, essas seções fornecem a base teórica necessária para análise e interpretação dos resultados e metodologias apresentadas ao longo dos demais capítulos.

2.1 Escoamento bifásico gás-líquido

O escoamento multifásico ocorre quando mais de uma fase escoar simultaneamente em um sistema. Um caso particular muito comum na produção de petróleo e gás é o escoamento bifásico líquido-gás, que dependendo das condições de operação e geometria do sistema de escoamento, pode apresentar diferentes distribuições geométricas entre as fases. Nesta seção, serão abordadas as principais características do escoamento bifásico gás-líquido, destacando os padrões de escoamento, os parâmetros, os mapas de padrões de escoamento e modelos propostos na literatura.

2.1.1 Padrões de escoamento

Em escoamentos bifásicos, as fases podem se organizar segundo diferentes configurações geométricas ao longo dos dutos. Essas configurações podem ser organizadas em grupos semelhantes, denominados padrões de escoamento, caracterizados pela distribuição espacial das interfaces entre as fases. Devido à natureza subjetiva com que esses padrões de escoamento são caracterizados, existem nomes alternativos e diferentes subdivisões para o mesmo padrão de escoamento. A identificação desses padrões de escoamento teve início com a visualização de imagens de tubos transparentes, posteriormente métodos mais objetivos foram desenvolvidos, como uso de sensores e análise de funções de probabilidade de parâmetros que caracterizam o escoamento (Bertola, 2003).

O padrão de escoamento em um sistema bifásico depende de parâmetros operacionais, variáveis geométricas e propriedades físicas das duas fases (Shoham, 2006). Um dos fatores determinantes é o ângulo de inclinação, a grande maioria dos estudos levam em conta pequenos ângulos em relação a tubulações verticais e horizontais, e alguns consideram a faixa completa de ângulos. Um dos principais estudos que propôs uma classificação sistemática dos padrões de escoamento foi apresentado por Barnea, Shoham e Taitel (1982). Os padrões definidos abrangem tubos horizontais e verticais, além da adição de pequenos ângulos em relação a essas referências.

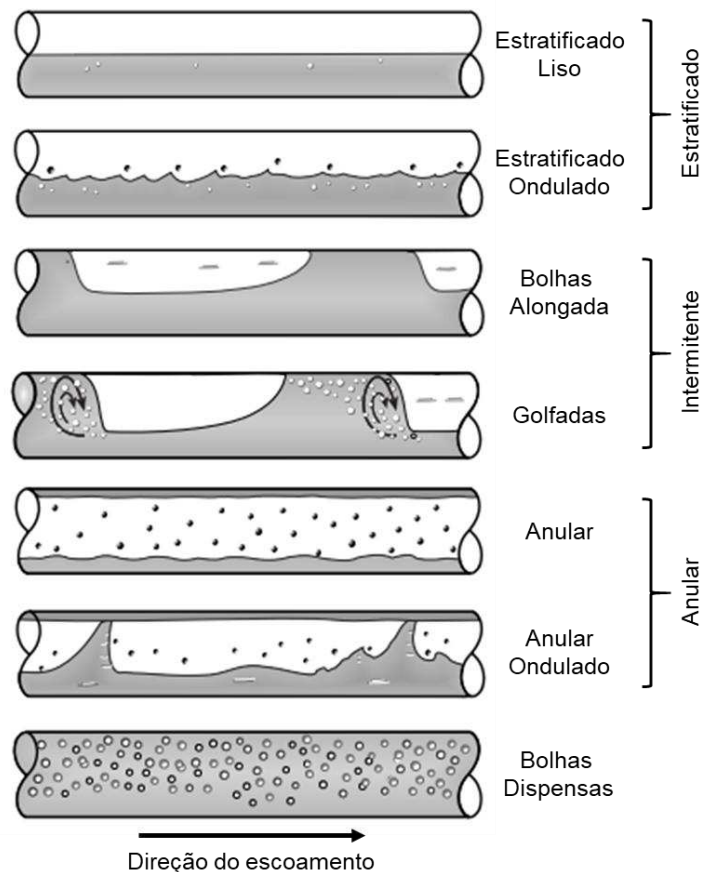
Em tubulações horizontais, a variação da pressão é geralmente mais suave, e a velocidade das fases tende a ser mais homogêneas. Segundo Shoham (2006), os padrões de escoamento em tubos horizontais e levemente inclinados, apresentados na Figura 1, podem ser divididos em quatro tipos principais, com subdivisões adicionais:

- **Estratificado:** Esse padrão de escoamento geralmente ocorre em condições de baixas vazões de líquido e gás. Devido ação da gravidade, a interface entre as fases é bem definida, com o líquido escoando na parte inferior e o gás na parte superior. O escoamento estratificado é dividido em duas classificações: liso ou ondulado. Essas classificações estão relacionadas ao formato da interface entre as fases, que podem ser onduladas dependendo da velocidade do gás.
- **Intermitente:** Trata-se de um regime onde as fases escoam alternadamente ao longo da tubulação, caracterizado pela presença de bolsões de gás envoltos por um filme líquido, intercalados com regiões de líquido puro. O escoamento intermitente pode ser classificado em bolhas alongadas e golfadas. O escoamento em bolhas alongadas acontece a baixas vazões de gás, resultando em um escoamento mais estável com uma menor presença de bolhas pequenas na região de líquido. Já o escoamento em golfadas a velocidade do gás é maior, com maior interação entre as fases.
- **Anular:** É um padrão de escoamento que acontece em altas vazões de gás. A fase gasosa escoar no centro da seção transversal da tubulação, enquanto a fase líquida escoar em um filme ao redor da parede do tubo. No caso do escoamento horizontal, geralmente, o filme da parte inferior é mais

espesso. Quando um escoamento anular apresenta uma interface altamente instável e uma quantidade alta de ondas aeradas, podemos defini-lo como Anular Ondulado.

- Bolhas Dispersas: É considerado um padrão de escoamento homogêneo e sem deslizamento. A altas vazões de líquido, as bolhas de gás são dispersas uniformemente na fase líquida e ambas as fases possuem a mesma velocidade.

Figura 1 – Padrões de escoamento em tubulações horizontais e quase horizontais

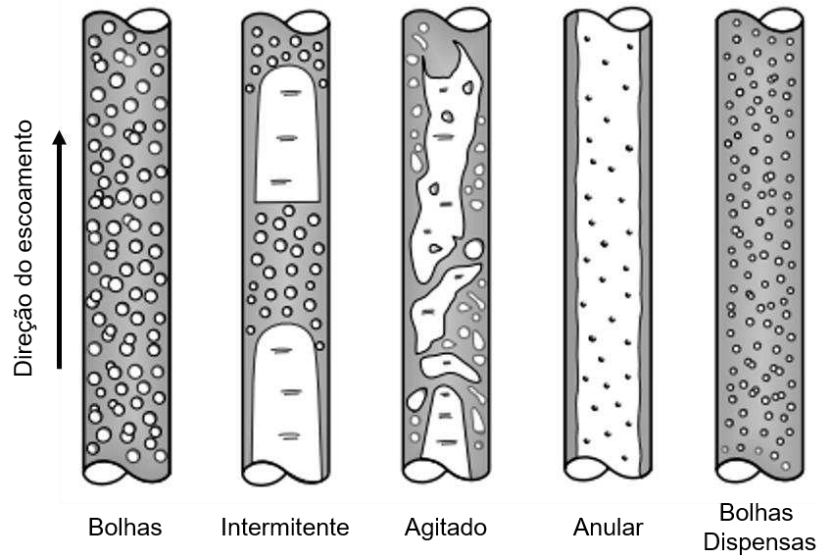


Fonte: Adaptado de Shoham (2006, p. 11).

No caso de tubulações verticais e fortemente inclinadas, devido a influência da gravidade, o padrão de escoamento estratificado não se desenvolve. Normalmente, esses padrões são mais simétricos em relação ao eixo central da tubulação. Segundo Shoham (2006), os padrões de escoamento apresentados na Figura 2, podem ser divididos em cinco classificações:

- Bolhas: É um padrão caracterizado por um alto deslizamento entre as fases líquidas e gasosas, devido a diferença entre a densidade do líquido e do gás. A fase gasosa é composta por diversas bolhas que se deslocam para cima, em um movimento sinuoso. As bolhas são aproximadamente homogêneas em relação a seção transversal da tubulação e a vazão da fase líquida é baixa.
- Intermitente: Este padrão apresenta, alternadamente, bolsões de gás com envoltos de filmes de líquido, seguidos de regiões de líquido aerado. Diferentemente do caso horizontal, as bolhas são simétricas em relação ao eixo da tubulação e possuem um diâmetro aproximadamente igual ao da tubulação.
- Anular: É um escoamento formado por uma fase gasosa com gotículas da fase líquida no centro da seção transversal da tubulação, com um filme líquido em contato com a parede da tubulação. O filme líquido nas paredes tem uma espessura uniforme e a interface entre as fases é ondulada.
- Agitado: Consiste em um padrão de escoamento semelhante ao intermitente, contudo, as regiões de líquido e gás apresentam uma delimitação menos nítida. O comportamento das regiões de líquido e gás são caóticos, fazendo com que diferentes bolsões de gás se fundam e que os pistões de líquido se tornem mais aerados.
- Bolhas Dispersas: De forma similar as tubulações horizontais, esse padrão de escoamento apresenta uma fase líquida a uma alta vazão, que carrega pequenas bolha de gás de forma homogênea e sem deslizamento.

Figura 2 – Padrões de escoamento em tubulações verticais e fortemente inclinadas



Fonte: Adaptado de Shoham (2006, p. 12).

Os mesmos padrões de escoamento bifásico gás-líquido ocorrem no caso ascendente e descendente. A diferença é que os mesmos padrões ocorrem em faixas diferentes de vazões das fases, por exemplo, baixas vazões de gás e líquido, podem gerar um escoamento de bolhas no caso ascendente e um escoamento anular no caso descendente (Bertola, 2003).

2.1.2 Parâmetros de escoamento

Diferentemente do escoamento monofásico, que pode ser caracterizado por um conjunto pequeno de parâmetros, o escoamento bifásico demanda a definição de mais parâmetros devido às interações entre as fases. É necessário considerar as propriedades intrínsecas das duas fases, como densidade, viscosidade, compressibilidade e difusividade. Essas propriedades definem como os fluidos interagem com as superfícies e forças externas, como a pressão. Os parâmetros do escoamento caracterizam o comportamento dos fluidos em movimento em um sistema de escoamento. Esses parâmetros dependem das características geométricas do sistema, das propriedades dos fluidos e da forma como eles se movem em relação ao tempo. A vazão mássica (W) e a vazão volumétrica (Q), podem ser definidos, respectivamente, como as quantidades de massa e volume de fluido que atravessam a seção transversal por unidade de tempo (Shoham, 2006).

No caso de escoamento de fluidos em tubulações, a velocidade não é constante em toda seção transversal. O perfil de velocidade descreve como a

velocidade varia na seção transversal, em uma tubulação, a velocidade é maior no centro e diminui até zero nas paredes devido à condição de não-escorregamento. A velocidade média pode ser definida como a média das velocidades em uma seção transversal. Isso é conveniente porque simplifica a análise do escoamento, facilitando cálculos devido ao fato que, diferentemente dos perfis de velocidade, a velocidade é um valor que permanece constante se a área da seção transversal for constante (Çengel; Cimbala, 2012). Para o escoamento bifásico, a velocidade superficial (J) de uma fase é definida como a razão entre o fluxo volumétrico dessa fase e a área total da seção transversal (A) do escoamento. Logo, em um escoamento bifásico gás-líquido, a velocidade superficial do gás (J_G) e a velocidade superficial do líquido (J_L) em uma seção transversal de uma tubulação podem ser definidas pelas Equações 1 e 2.

$$J_L = \frac{Q_L}{A} \quad (1)$$

$$J_G = \frac{Q_G}{A} \quad (2)$$

A velocidade superficial da mistura (J) pode ser calculada pela razão entre a soma das vazões volumétricas e a área total da seção transversal. Outra forma é somar as velocidades superficiais da fase líquida e da fase gasosa, como dado pela seguinte expressão:

$$J = J_G + J_L = \frac{Q_G + Q_L}{A} \quad (3)$$

Além das velocidades das fases, as propriedades físicas das fases também influenciam significativamente no comportamento do escoamento bifásico. Entre essas propriedades destacam-se a densidade da fase gasosa (ρ_G) e a densidade da fase líquida (ρ_L). Adicionalmente, identificar a distribuição espacial das fases ao longo da tubulação é fundamental para caracterização do escoamento, uma vez que influencia diretamente o padrão de escoamento, a perda de carga e a dinâmica das fases. A fração de líquido (H_L), também conhecido como *holdup*, é a fração do volume total de uma região específica do escoamento que é ocupada por líquido. Já a fração de gás (α), mais conhecida como fração de vazio, é a fração complementar e refere-se à fração de gás que ocupa esta região. Esses conceitos são complementares, ou

seja, cada fração pode assumir valores entre 0 e 1, e a sua soma é sempre igual a 1. A fração de gás e a fração de líquido, podem ser relacionadas pela seguinte equação:

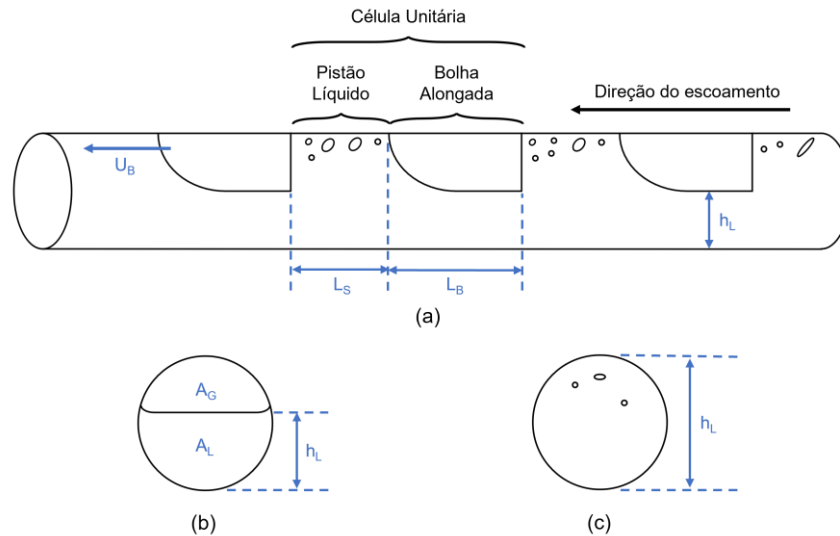
$$\alpha = 1 - H_L \quad (4)$$

O escoamento bifásico intermitente apresenta um comportamento repetitivo, com uma sequência de estruturas compostas bolhas alongadas envoltas por um filme de líquido e pistões de líquido aerados. Essas estruturas repetitivas, foram inicialmente discutidas e sistematizadas por Wallis (1969), para a análise do escoamento vertical intermitente. Essas estruturas são denominadas células unitárias, nos modelos de escoamentos baseados nessa abordagem, cada célula é definida como a região do escoamento intermitente entre os narizes de duas bolhas alongadas consecutivas. A Figura 3 apresenta uma representação de uma tubulação horizontal com escoamento intermitente, indicando uma célula unitária, os parâmetros que a caracterizam e as seções transversais do escoamento na região de bolha alongada e pistão de líquido.

As células unitárias são caracterizadas através de alguns parâmetros fundamentais, distintos para região da bolha alongada e do pistão de líquido. A velocidade de translação da bola (U_B) refere-se à velocidade com que a bolha alongada de uma célula unitária se desloca ao longo da tubulação. O comprimento de pistão líquido (L_S) e o comprimento de bolha alongada (L_B), definem o tamanho da célula unitária, com uma divisão no final da bolha alongada. A altura de líquido (h_L) é uma medida que define a distribuição da fase líquida dentro da tubulação. Ela representa distância entre a interface entre as fases e a tubulação. E por fim, a frequência (f) é definida como o inverso do período de passagem de uma célula unitária por uma seção transversal.

Essa representação permite analisar o escoamento de forma modular, ou seja, a análise complexa do escoamento intermitente é simplificada quebrando-a em partes menores com comportamento mais repetitivo e previsível. Essa divisão facilita a aplicação de modelagem matemática e numérica, dado que o comportamento global do escoamento pode ser descrito através do conjunto das células unitárias e suas interações.

Figura 3 – Representação da célula unitária: (a) Célula unitária; (b) Seção transversal da bolha alongada; (c) Seção transversal do pistão de líquido



Fonte: Autoria própria.

Diferentes modelos e correlações presentes na literatura utilizam parâmetros e números adimensionais para descrever o comportamento do escoamento intermitente, permitindo uma análise válida para diferentes geometrias e tipos de fluídos. Entre os parâmetros adimensionais mais utilizados destaca-se o número de Froude, que pode ser definido como a razão entre as forças de inércia e a força gravitacional, sendo expresso pela Equação 5.

$$Fr = \frac{J}{\sqrt{gD}} \quad (5)$$

Onde:

- J : Velocidade superficial da mistura;
- g : aceleração da gravidade;
- D : diâmetro da tubulação.

No contexto das bolhas alongadas, a velocidade pode ser representada por uma relação linear, conforme proposto por Nicklin (1962), que expressa a velocidade da bolha em função da velocidade superficial da mistura. Essa relação é apresentada na Equação 6.

$$U_B = C_0 J + u_\infty \quad (6)$$

Onde:

- U_B : Velocidade da bolha alongada;
- C_0 : parâmetro de distribuição;
- u_∞ : velocidade de deriva.

Se dividimos os dois lados da equação $\sqrt{gD \left(\frac{\rho_L - \rho_G}{\rho_L} \right)}$, obtemos as Equações 7 e 8. As equações estabelecem uma relação linear entre o número de Froude e a velocidade da bolha alongada.

$$\frac{U_B}{\sqrt{gD \left(\frac{\rho_L - \rho_G}{\rho_L} \right)}} = C_0 \frac{J}{\sqrt{gD \left(\frac{\rho_L - \rho_G}{\rho_L} \right)}} + \frac{u_\infty}{\sqrt{gD \left(\frac{\rho_L - \rho_G}{\rho_L} \right)}} \quad (7)$$

$$\frac{U_B}{\sqrt{gD \left(\frac{\rho_L - \rho_G}{\rho_L} \right)}} = C_0 F_r + C_\infty \quad (8)$$

Outro número adimensional frequentemente utilizado é o número de Strouhal, que relaciona a frequência de passagem das bolhas alongadas ao diâmetro do tubo e à velocidade superficial do gás. Esse parâmetro, definido na Equação 9, é amplamente empregado em estudos que comparam diferentes condições experimentais, como em Fossa (2003).

$$S_t = \frac{fD}{J} \quad (9)$$

Onde:

- S_t : Número de Strouhal;
- f : frequência;

Outros parâmetros importantes são as frações de líquido e gás, cujo cálculo pode variar conforme a metodologia adotada. Dependendo da região de interesse, essas frações podem ser definidas considerando a célula unitária como um todo ou

separando as regiões de pistão de líquido e de bolha alongada. Outra abordagem comum é do cálculo de frações de fase em cada seção transversal, encontrando a proporção das áreas correspondentes as fases e a área total da seção transversal da tubulação. Na região do pistão de líquido, a fração de líquido (H_{LS}), é a razão entre o volume ocupado por líquido (V_{LS}) e o volume total da região do pistão de líquido (V_S). Já na região da bolha alongada, a fração de líquido (H_{LB}) corresponde a razão entre o volume ocupado por líquido (V_{LB}) e o volume total da região da bolha alongada (V_B). As frações de gás para cada região podem ser calculadas com base nas frações de líquido, dado que são parâmetros complementares. As Equações 10 e 11 apresentam os cálculos descritos.

$$H_{LS} = \frac{V_{LS}}{V_S} \quad (10)$$

$$H_{LB} = \frac{V_{LB}}{V_B} \quad (11)$$

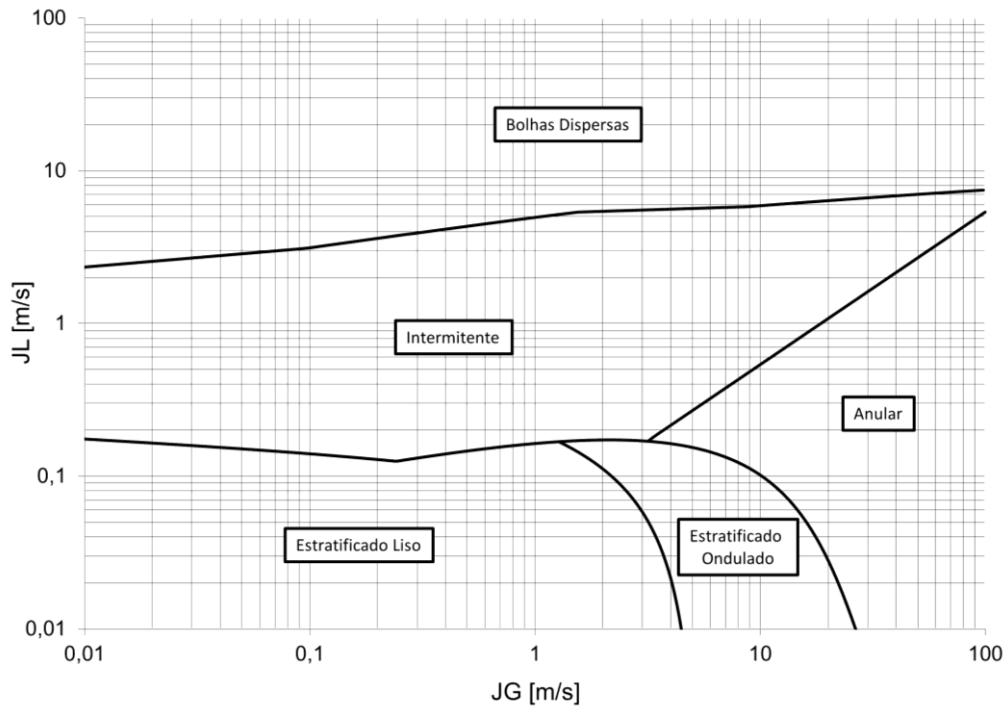
O formato das bolhas alongada é uma informação que também pode ser explorada. O estudo de Pipa et al. (2014) destaca a importância da caracterização do formato de bolhas em um escoamento bifásico líquido-gás, propondo uma abordagem baseada em processamento de imagens e técnicas de problema inverso. Diferentes condições de escoamento podem gerar diferentes formatos de nariz e cauda das bolhas alongadas. A forma das bolhas impacta diretamente na dinâmica do escoamento, uma análise profunda possibilita a correção de modelos de predição e de modelagem computacional.

2.1.3 Mapa de padrões de escoamento

Um mapa de padrão de escoamento é uma representação gráfica bidimensional que classifica diferentes padrões de escoamento com respaldo em dois parâmetros que caracterizam o escoamento. Os primeiros mapas foram propostos na década de 50, com base em dados experimentais. Um dos estudos mais relevantes para escoamentos em tubulações horizontais com escoamento bifásico gás-líquido foi elaborado por Taitel e Dukler (1976), com base na elaboração de um modelo teórico e na validação deste com dados experimentais. Além disso, o estudo analisou a transição entre os padrões de escoamento e a influência de forças físicas, para tubulações horizontais e levemente inclinadas. O mapa é apresentado na Figura 4,

relaciona a velocidade superficial do líquido no eixo x e a velocidade superficial do gás no eixo y .

Figura 4 – Mapa de escoamento bifásico líquido-gás



Fonte: Adaptado de Taitel e Dukler (1976).

2.2 Processamento de imagens

Este capítulo apresenta conceitos relacionados a representação de uma imagem digital e transformações (por exemplo, binarização e filtragem espacial). Uma revisão mais aprofundada é elaborada abordando técnicas de segmentação de imagem e extração de características. Ao final, algumas técnicas para visualização e análise de escoamento são apresentadas, apresentando estudos que adotam processamento de imagens de alta velocidade para extração de parâmetros do escoamento.

2.2.1 Conceitos básicos

O processamento digital de imagens refere-se ao conjunto de técnicas aplicadas a imagens digitais para manipular, analisar e transformar imagens digitais. Uma imagem digital pode ser definida como uma função bidimensional $f(x, y)$, onde f é uma amplitude, finita, positiva e discreta, e está associada a cada par de coordenadas (x, y) . As imagens digitais são compostas por elementos discretos, cada

um com uma posição específica e um valor de intensidade. Esses elementos são denominados pixels, e representam a menor unidade de uma imagem digital.

O valor f para um determinado par de coordenados é denominado intensidade. Ele pode ser caracterizado pela quantidade de iluminação que incide sobre a cena está sendo vista e a quantidade de iluminação que é refletida sobre os objetos na cena (Gonzalez; Woods, 2009). A quantidade de iluminação é denominada componente de iluminação e está associada à energia irradiada pela fonte. Já a fração refletida pelos objetos corresponde à componente de refletância, que pode variar de 0 (absorção total) a 1 (reflexão total). O produto dos dois componentes caracteriza a intensidade. Logo, a intensidade pode ser dada por:

$$f(x, y) = i(x, y) \cdot r(x, y) \quad (12)$$

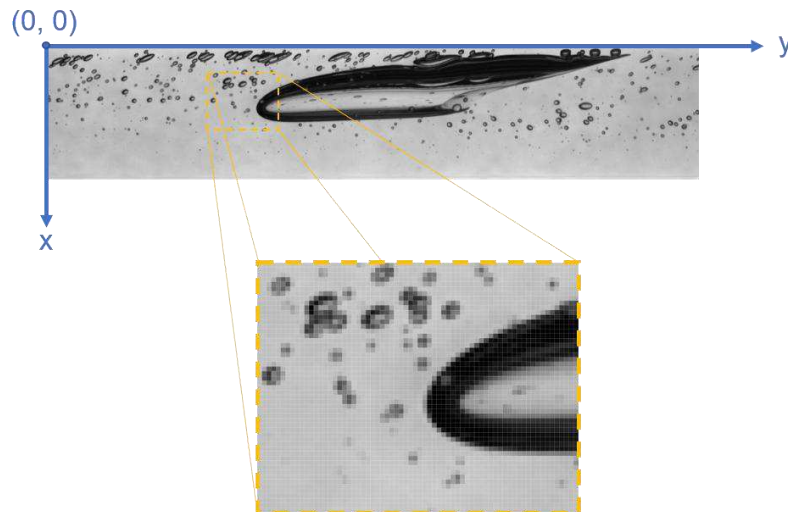
Para que uma imagem digital seja gerada, é necessário sensores capturem a energia proveniente da fonte de iluminação e a transformem em um sinal elétrico. Posteriormente, esse sinal é digitalizado e transformado em luz visível. Esses sensores, geralmente são dispostos em três tipos de arranjo: sensores de varredura única (*Single Sensor*), sensores lineares (*Line Sensor Array*) e sensores matriciais (*Area Sensor Array*). A escolha do tipo de arranjo e do material sensor depende da aplicação, variando conforme os requisitos técnicos específicos de cada caso e da fonte de energia analisada (Gonzalez; Woods, 2009).

Os sensores matriciais são predominantemente utilizados nas câmeras digitais, com destaque para o sensor do tipo CMOS (*Complementary Metal-Oxide-Semiconductor*). Os sensores CMOS são amplamente utilizados em aplicações que exigem captura rápida de imagens. Um arranjo matricial típico desses sensores é de 16 megapixels (4000 x 4000 pixels), onde a imagem completa é capturada de forma instantânea (*global shutter*). Outras propriedades significativas dos sensores CMOS são o baixo custo de produção e o baixo consumo energético.

Na grande maioria das aplicações, as imagens são representadas através de matrizes numéricas 2D. No caso de imagens monocromáticas, apenas uma matriz é utilizada. Cada pixel é representado por um valor numérico, que varia com diferentes tonalidades de uma única cor. Para imagens coloridas, a representação é dada através da combinação três ou quatro matrizes numéricas 2D (denominadas canais). Essas combinações dependem do padrão de representação de cor escolhido, os padrões mais comuns são baseados vermelho, verde e azul (RGB). Neste caso, cada

pixel depende de uma combinação entre os valores dos elementos de cada canal utilizado. A Figura 5 apresenta a representação visual de uma bolha alongada em imagem digital monocromática digital em escala de cinza, cada pixel é representado em tons de cinza, variando de 0 (preto) até 255 (branco).

Figura 5 – Imagem em escala de cinza de uma tubulação horizontal em um sistema de escoamento bifásico



Fonte: Autoria própria.

As medidas de resolução de imagens digitais definem o nível de detalhe que uma imagem pode representar. A ampliação na Figura 5 demonstra que a representação visual de uma imagem é composta por um conjunto de pixels. A resolução espacial de uma imagem digital pode ser representada pelo número de pixels na largura e altura da imagem. Já a resolução de intensidade é representada pelo número de bits por pixel, ela define o número de níveis de intensidade de cada canal. O exemplo apresentado na figura possui uma resolução espacial de 828x148 pixels e uma resolução de intensidade de 8 bits (255 níveis de cinza).

2.2.2 Segmentação de imagens

A segmentação de imagens é o processo de particionar uma imagem digital em regiões distintas, agrupando um conjunto de pixels com base em algum tipo de similaridade. O propósito específico da segmentação depende de cada aplicação, mas na maioria dos casos está associado a redução da complexidade, extração de informações relevantes e detecção de objetos. Esse processo é essencial em diversas

áreas que utilizam análise de imagens digitais, como medicina, robótica, segurança e mecânica dos fluídos.

É possível categorizar a segmentação de imagens digitais em três tipos principais: Segmentação Semântica, Segmentação de Instância e Segmentação Panóptica (Brilhador; Lazzaretti; Lopes, 2025). Na segmentação semântica, cada pixel é rotulado em categorias predefinidas. Já a segmentação de instâncias, além de categorizar cada pixel da imagem, distingue os objetos individuais pertencentes a uma mesma categoria. E por fim, a segmentação panóptica é uma combinação dos outros dois tipos de segmentação, classificando os pixels diversas categorias e identificando os diversos objetos dentro de cada categoria.

Inicialmente, as técnicas de segmentação de imagens eram baseadas exclusivamente em métodos determinísticos e heurísticos, como processos morfológicos, aplicação de filtros e estratégias fundamentadas em regras predefinidas. Com o avanço da tecnologia, técnicas baseadas em aprendizado de máquina foram desenvolvidas para segmentação de imagens, aprimorando as abordagens tradicionais. Os tópicos a seguir apresentam as principais técnicas de segmentação de imagens, assim como métodos auxiliares que aprimoram esses processos.

2.2.2.1 Técnicas baseadas em limiarização

As técnicas mais simples para segmentação incluem transformadas de intensidade no domínio espacial. Essas técnicas se resumem em funções que mapeiam os valores de cada pixel da imagem original para um novo valor na imagem resultante, dependendo apenas da intensidade em cada pixel. A limiarização (*thresholding*) é uma técnica baseada em transformadas de intensidade, onde todos os valores de intensidade são reduzidos em dois ou mais níveis. Para executar esta separação, um limiar de intensidade é definido. O nível inferior geralmente representado pelo valor 0, correspondendo à cor preta, enquanto o nível superior é representado como o valor máximo, correspondendo à cor branca, que pode variar dependendo da profundidade de bits da imagem.

Várias técnicas podem ser empregadas para aplicação de limiarização, essas abordagens variam em relação ao número de níveis de cor, ao cálculo e automação dos limiares, ao número de limiares e à adaptação a variações locais. Uma técnica comum e simples é o limiar global, onde um limiar de intensidade é selecionado para

todos os pixels da imagem. Um exemplo de função de limiarização global é apresentado na Equação 13.

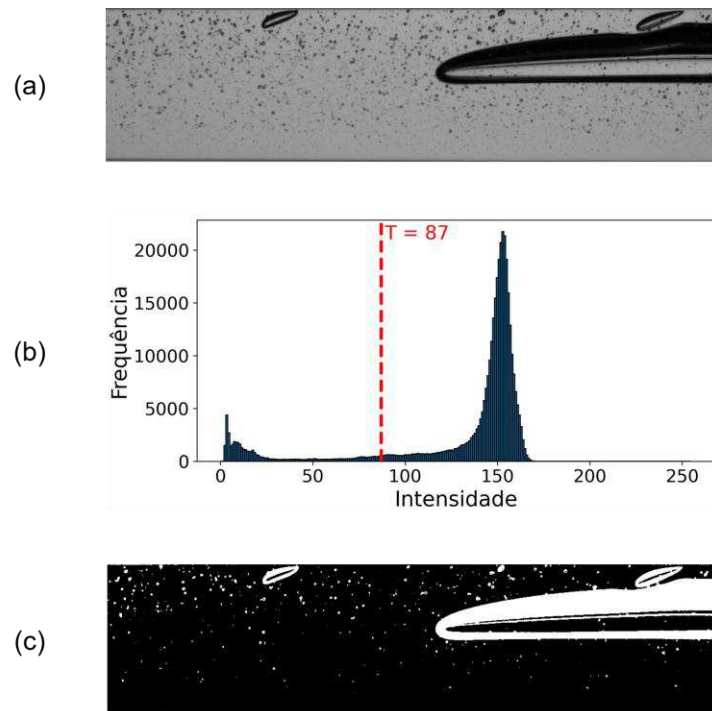
$$g(x,y) = \begin{cases} 0, & \text{se } f(x,y) < T \\ 255, & \text{se } f(x,y) > T \end{cases} \quad (13)$$

Onde:

- $f(x,y)$ representa a intensidade do pixel na posição (x,y) na imagem original;
- $g(x,y)$ representa a intensidade do pixel na posição (x,y) na imagem resultante;
- T é o limiar de intensidade escolhido.

A Figura 6 apresenta um exemplo do processo de segmentação baseado em limiar de intensidade. A subfigura (a) mostra a imagem em escala de cinza correspondente à região do nariz da bolha. A subfigura (b) apresenta o histograma de intensidade, no qual a linha vermelha tracejada indica o limiar ($T = 87$) adotado para obter a imagem binária apresentada na subfigura (c).

Figura 6 – Exemplo de segmentação com limiar de intensidade



Fonte: Autoria própria.

Encontrar um limiar global muitas vezes pode ser uma tarefa desafiadora. Para sanar tal problema, algumas metodologias estipularam formas de definir um

limiar de forma automática, com base em técnicas estatísticas. A solução proposta por Otsu (1979), se baseia em dividir o histograma da imagem em duas classes (objeto e fundo da imagem). O limiar é definido através da análise da variância da intensidade, minimizando a variância dentro de cada classe e maximizando a variância entre as classes. Ou seja, os valores de intensidade dentro de uma classe devem ser o mais homogêneos possível e a média das intensidades das classes deve ser o mais distante possível. A aplicação do método automatiza e padroniza a escolha do limiar de intensidade, evitando a falta de consistência de métodos manuais e heurísticos.

As imagens digitais podem apresentar variações de iluminação, ruídos e contraste não uniforme, fazendo com que um único limiar não seja suficiente para segmentar a imagem corretamente. Algumas técnicas utilizam limiares adaptativos, onde o valor do limiar varia localmente, de acordo com as características de diferentes regiões da imagem. Um exemplo de metodologia foi proposta por Sauvola e Pietikak (2000), com o objetivo de segmentar imagens de documentos. Os limiares para cada região foram definidos com base na média local e desvio da intensidade.

Outra característica importante é o número de níveis de intensidade da imagem resultante de segmentação. Em alguns casos, é necessário dividir uma imagem em mais de duas classes, com base em múltiplos limiares. Por exemplo, em imagens médicas, diferentes tecidos podem apresentar intensidades distintas, possibilitando a identificação individual.

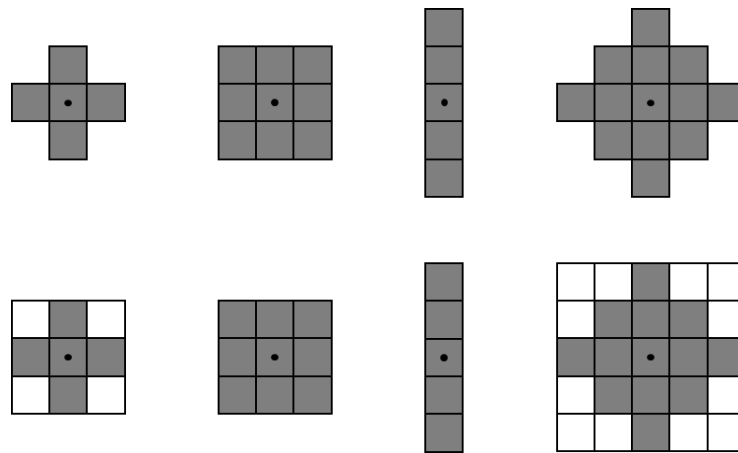
2.2.2.2 Técnicas baseadas em morfologia

A morfologia matemática é uma abordagem para análise de estruturas espaciais baseadas em teoria dos conjuntos, formalizada e sistematizada por Serra (1984), dando origem as primeiras operações morfológicas para análise e processamento de imagens. Os objetos presentes nas imagens digitais são representados como conjuntos em um espaço discreto. Através da exploração da forma e estrutura desses objetos é possível extrair características, remover ruídos e segmentar regiões de interesse.

Duas entidades fundamentais compõe a morfologia matemática: a imagem digital e o elemento estruturante. Em uma imagem digital, cada pixel é tratado como um elemento de um conjunto. No caso de imagens em escala de cinza, sua representação pode ser dada por conjuntos com componentes tridimensionais, com dois componentes espaciais e um de intensidade. Os elementos estruturantes são

pequenos conjuntos utilizados para modificar ou analisar uma imagem, definindo como cada pixel será tratado com base nos seus vizinhos (Gonzalez; Woods, 2009). A Figura 6 apresenta alguns exemplos de elementos estruturantes. Os quadrados cinzas representam os membros do elemento estruturante específico. A origem de cada elemento estruturante é indicada pela presença de um ponto preto no centro de um membro. Para que sejam utilizados com imagens, os elementos estruturantes da primeira linha da Figura 7 precisam transformados em arranjos matriciais retangulares. Para isso, a quantidade mínima necessária de elemento é adicionada ao fundo (quadrados brancos).

Figura 7 – Elementos estruturantes e conversão em arranjos retangulares



Fonte: Adaptado de Gonzalez e Woods (2009).

Para imagens digitais, os arranjos retangulares são representados como matrizes numéricas. Geralmente, essas matrizes são chamadas de elementos estruturantes ou kernels morfológicos. Para imagens binárias, as essas matrizes recebem o valor de 1 para as posições dos elementos estruturante e 0 para posições adicionadas para completar a matriz retangular. No caso de imagens em escala de cinza ou coloridas, essas matrizes podem receber valores diferentes de 0 e 1. Essas matrizes são utilizadas pelas operações morfológicas para modificar a forma dos objetos presentes na imagem. As operações morfológicas combinam as informações de intensidade na vizinhança de um pixel, para definir como cada pixel será alterado na imagem resultante. Essas operações podem expandir ou reduzir objeto em uma imagem, as operações fundamentais para imagens binárias são:

- **Erosão:** Para cada posição, verifica quais pixels da imagem são cobertos pelos pixels do elemento estruturante. Se todos os pixels cobertos forem 1,

o pixel central também será 1 na imagem resultante. Caso contrário, o pixel central será substituído por 0 na imagem resultante. Essa operação é utilizada para reduzir o tamanho dos objetos na imagem binária, e em alguns casos, os detalhes menores são removidos.

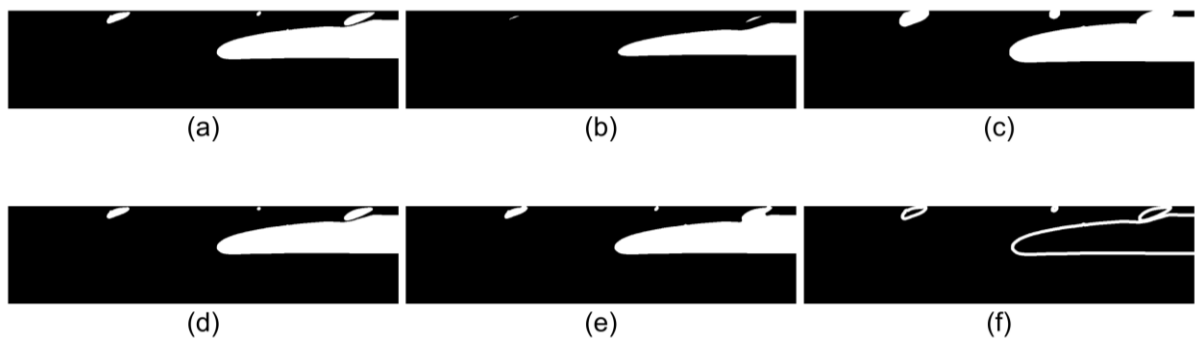
- **Dilatação:** Para cada posição, verifica quais pixels da imagem são cobertos pelos pixels do elemento estruturante. Se todos os pixels cobertos forem 0, o pixel central também será 0 na imagem resultante. Caso contrário, o pixel central será substituído por 1 na imagem resultante. Diferentemente da erosão, essa operação é utilizada para aumentar o tamanho dos objetos na imagem binária, mudando o formato e espessando os objetos brancos de acordo com o tamanho e formato do elemento estruturante.
- **Abertura:** É uma operação resultante de uma erosão seguida de uma dilatação. O resultado da sua aplicação é a suavização das bordas, a remoção de pequenos detalhes e separação de objetos conectados.
- **Fechamento:** É uma operação resultante de uma dilatação seguida de uma erosão. Assim como a abertura, também suaviza as bordas dos objetos. Diferentemente da abertura, tende a fechar buracos dentro de objetos, conectar objetos próximos e destacar detalhes pequenos.

Algumas operações matemáticas podem ser aplicadas diretamente em cada pixel utilizando imagens digitais. Outra abordagem possível é combinar essas operações matemáticas e as operações morfológicas fundamentais apresentadas anteriormente (erosão, dilatação, abertura e fechamento). Alguns exemplos disso são:

- **Gradiente Morfológico:** é definido como a diferença entre a erosão e a dilatação de uma imagem digital. É uma operação aplicada para destacar as bordas e transições de intensidade.
- **Top-hat (white Top-hat):** é definida como a diferença entre uma imagem e a sua abertura. Essa operação morfológica realça regiões claras menores que o elemento estruturante.
- **Bottom-hat (black Top-hat):** é definida como a diferença entre o fechamento morfológico de uma imagem e a própria imagem.

Alguns efeitos dessas operações incluem o fechamento de buracos, a eliminação de ruídos, o preenchimento de regiões desconectadas, o espessamento, o afinamento e a separação de objetos conectados. Esses efeitos podem ser utilizados complementar técnicas de segmentação de imagens, tanto no pré-processamento ou quanto no pós-processamento. A Figura 8 apresenta a aplicação das operações morfológicas à imagem do nariz de uma bolha alongada segmentada manualmente, sendo mostradas em (a) a imagem binária original, (b) a erosão, (c) a dilatação, (d) a abertura, (e) o fechamento e (f) o gradiente morfológico.

Figura 8 – Exemplos de operações morfológicas em máscara binária



Fonte: Autoria própria.

2.2.2.3 Técnicas baseadas em regiões

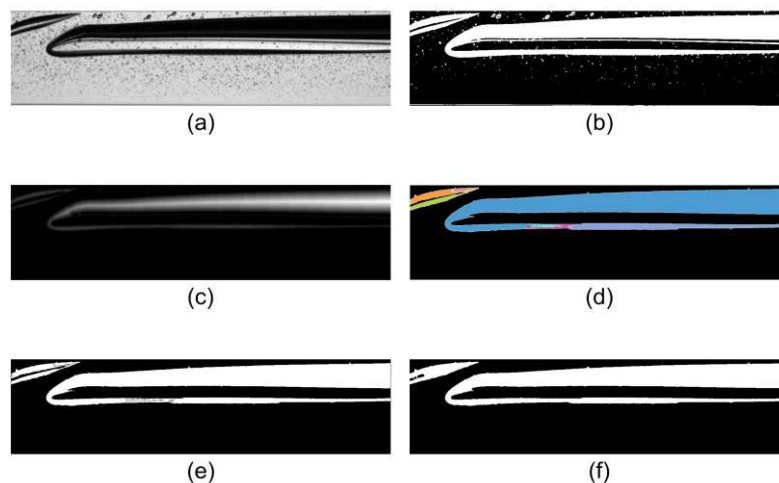
As técnicas de segmentação baseadas em regiões buscam agrupar pixels com base em propriedades predefinidas, como cor, intensidade, textura ou momentos estatísticos. A partir de pixels iniciais chamados sementes, uma região homogênea é incrementada com os pixels vizinhos que satisfaçam os critérios de similaridade. A decisão dos pontos de partida, critérios de similaridade e regra de parada dependem tanto do tipo de aplicação quanto do tipo de imagem. O algoritmo *Seeded Region Growing* (SRG), proposto por Adams e Bischof (1994), é um exemplo de segmentação baseada em crescimento de região. O método requer apenas a seleção de um conjunto de pontos iniciais, sem a necessidade de parâmetros de ajuste adicionais. Posteriormente, os pixels vizinhos das sementes são adicionados seguindo um critério de crescimento e a região é atualizada. O procedimento é repetido até que todos os pixels da imagem sejam rotulados ou que nenhum pixel candidato satisfaça o critério de similaridade.

Outra abordagem possível para segmentação baseada em região é a divisão e fusão de região. Diferentemente do crescimento de região, inicialmente, a imagem

inteira é considerada uma região, dividindo-a em sub-regiões, que são fundidas, posteriormente, com base nos critérios de homogeneidade. Essa abordagem pode ser mais adaptável a variações locais, pois não depende de uma homogeneidade global. A metodologia também é mais robusta em relação a ruído, dado que sub-regiões ruidosas dificilmente são consideradas homogêneas, fazendo com que não sejam adicionadas a região de interesse.

O watershed, por sua vez, também é uma técnica de segmentação baseada em regiões. Essa metodologia tem como fundamento a visualização de uma imagem em duas coordenadas espaciais e uma coordenada de intensidade (Vicent, 1991). A imagem é vista como um mapa de altura, onde as regiões de baixa intensidade correspondem a vales e as regiões de alta intensidade correspondem a picos. A partir dos mínimos locais, diferentes regiões são preenchidas simultaneamente. Para evitar que essas regiões se fundam, contornos conhecidos como linhas de watershed são criados de modo que apenas eles fiquem visíveis no final do processo. Em resumo, o algoritmo de watershed funciona partindo de uma ideia de alagamento para segmentar as imagens, separando os objetos encontrando as fronteiras entre as regiões. A Figura 9 ilustra a aplicação do algoritmo watershed no processo de segmentação da subfigura (a). Observa-se que, após a binarização da imagem em (b) e o cálculo da transformada de distância em (c), o método gera múltiplas regiões rotuladas, conforme mostrado em (d). Em (e), apresenta-se a imagem segmentada após a etapa de rotulação. Já em (f), é mostrada a aplicação de operações morfológicas com objetivo de unir regiões que permaneceram separadas.

Figura 9 – Exemplo de aplicação do método watershed



Fonte: Autoria própria.

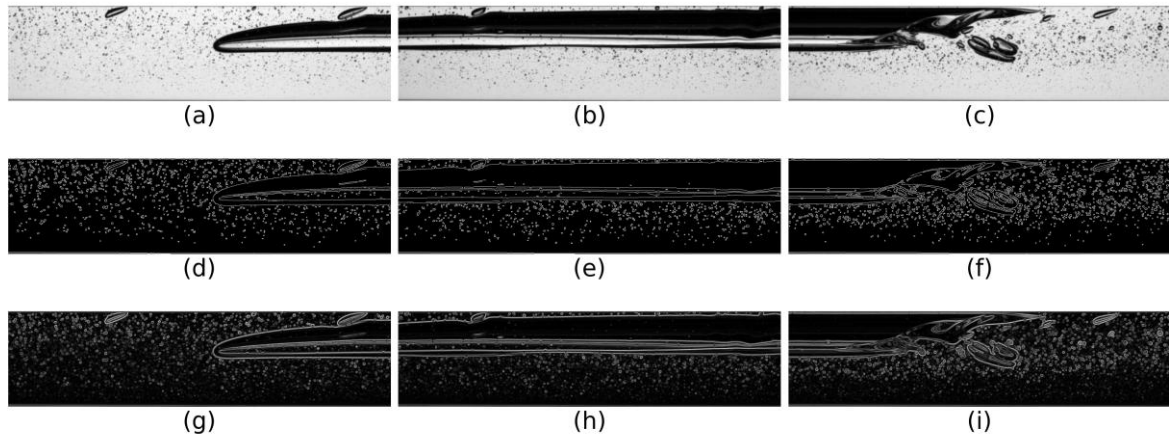
Há diversos algoritmos disponíveis para implementação da transformada de watershed. Algumas técnicas são implementadas para pré-processamento e para gerar informações auxiliares. Para imagens binárias, transformadas de distância podem ser aplicadas para separar objetos conectados e definir o centro dos objetos nas imagens. Já para imagens em escala de cinza, operadores de gradiente podem ser aplicados para destacar transições abruptas e obter informações sobre os limites das regiões a serem segmentadas. As abordagens citadas, podem causar uma divisão excessiva de regiões, principalmente na presença de irregularidades (Coelho, 2011). Uma metodologia que inibe esses problemas é a de utilização de marcadores de controle, que são previamente definidas e funcionam como pontos de partida para transformada de watershed. Os marcadores são imagens auxiliares que definem pontos de partida para regiões internas dos objetos (*foreground*) e regiões de fundo (*background*). Neste caso, a “inundação” da transformada watershed tem os marcadores como ponto de partida, definindo as bordas entre as regiões internas e fundo.

2.2.2.4 Técnicas baseadas em contornos

A segmentação baseada em contornos inclui técnicas com objetivo principal de encontrar os limites dos objetos por meio de discontinuidades em propriedades das imagens digitais, como intensidade, cor ou textura. Através dessas transições é possível encontrar a silhueta dos objetos e determinar a região pertencente a determinado objeto. Diferentes abordagens podem ser aplicadas para encontrar os contornos dos objetos, as técnicas mais simples se resumem em definir os pontos onde a intensidade da imagem muda abruptamente, com base nos pixels vizinhos. Para isso, operadores de borda como Sobel (Sobel, 1973) e Canny (Canny, 1986), definem e realçam os contornos por meio da aplicação de kernels sobre a imagem. Inicialmente, o gradiente das imagens é calculado, indicando a magnitude da variação de intensidade em cada ponto da imagem, na direção horizontal e vertical. Em seguida, as bordas são definidas com a aplicação de um limiar, os pixels pertencentes bordas das imagens são marcados como branco enquanto a ausência de borda é representada por pixels pretos. Alguns operadores como Canny, possuem etapas adicionais antes da aplicação do limiar, como suavização por meio de filtro Gaussiano e supressão de não-máximos. A Figura 10 apresenta exemplos da aplicação de operadores de detecção de bordas. As imagens (a-c) correspondem aos frames

originais em escala de cinza com diferentes regiões da bolha alongada. As imagens (d-f) mostram as bordas detectadas pelo algoritmo Canny, enquanto as imagens (g-i) apresentam o gradiente obtido pelo operador de Sobel.

Figura 10 – Detecção de bordas em imagens de escoamento bifásico utilizando os operadores Sobel e Canny



Fonte: Autoria própria.

Devido ao fato de serem baseados em gradiente ponto a ponto, os operadores de borda podem resultar em curvas irregulares e desconexas. Algumas metodologias, denominadas modelos deformáveis, segmentam objetos globalmente, utilizando curvas que se deformam dinamicamente aos contornos dos objetos na imagem, gerando contornos contínuos, fechados e suaves. Os modelos deformáveis também são conhecidos como Snakes, e foram inicialmente definidos por Kass e Witkin (1988). Inicialmente, uma curva é definida próxima ao contorno do objeto e a evolução do contorno é guiada pela minimização de uma energia total. Essa energia depende de duas forças, que estão relacionadas ao controle de suavidade da curva (força interna) e a uma atração da curva as bordas características da imagem (força externa). As forças internas evitam que os pontos da curva que está evoluindo se afastem e que dobras bruscas ocorram. Já a força externa, calculada através do gradiente, atrai a curva em evolução para regiões com variações bruscas de intensidade. As iterações são finalizadas quando a curva se estabiliza e atinge um valor mínimo de energia.

Embora os Snakes sejam eficazes para contornos suaves e definidos, eles podem apresentar limitações quando uma única curva precisa se dividir ou fundir com outras curvas. Para esses casos, a aplicação do método level set (Stanley Osher, 1988) pode ser uma solução mais robusta, principalmente para objetos com bordas contínuas e topologias complexas, com fusões e divisões. O level set é uma técnica

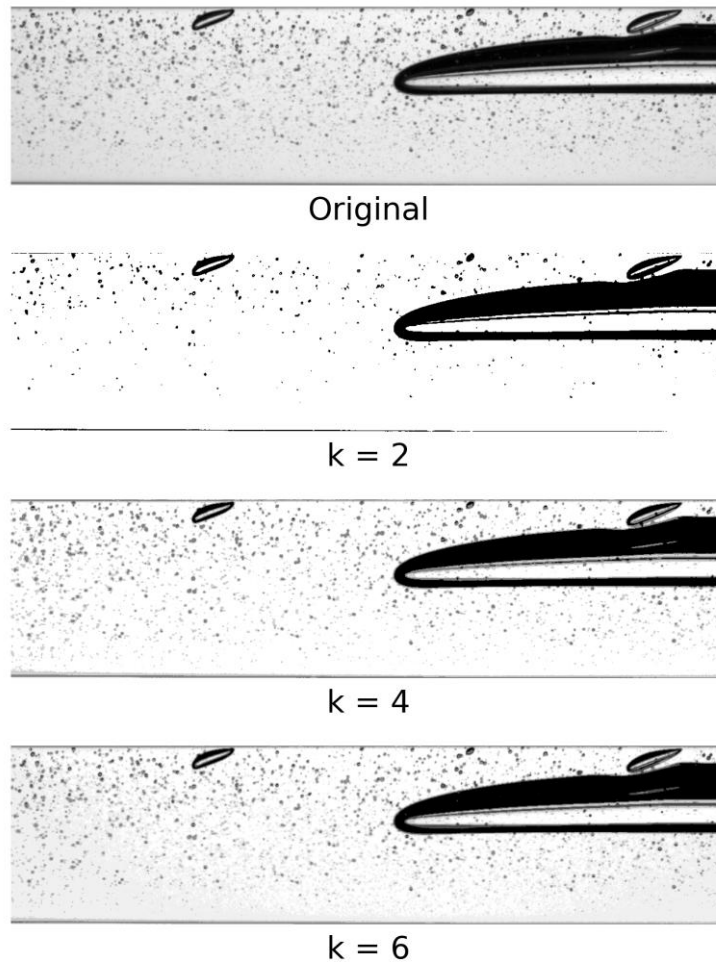
numérica baseada na equação de Hamilton-Jacobi, a ideia principal é representar uma interface como o conjunto dos valores iguais a zero de uma função escalar contínua (função level set). Inicialmente, o contorno pode ser definido manualmente ou por meio de alguma heurística. Dentro do contorno, a função level set assume valores negativos, enquanto fora do contorno seus valores são positivos. O contorno inicial evolui seguindo uma equação de evolução, que é guiada através das propriedades da imagem, como bordas, variações de intensidade e outras informações. Essa evolução é interrompida quando o contorno deixa de mudar significativamente ou um número de iterações máximo definido é atingido.

2.2.2.5 Técnicas baseadas em agrupamento

As técnicas de segmentação baseadas em agrupamento consistem em métodos que particionam os pixels da imagem em grupos de acordo com sua similaridade em atributos como intensidade, cor, textura ou posição espacial, sem necessidade de rótulos prévios, caracterizando-se como uma abordagem de aprendizado não supervisionado (Jain et al., 2000).

Em muitos algoritmos de agrupamento, como o K-Means (Macqueen, 1967), a segmentação depende da definição previa do número de grupos de k , que determinam quantos clusters serão formados a partir dos dados da imagem. Quanto maior o valor de k , maior o nível de detalhamento da segmentação, permitindo identificar pequenas variações nas imagens e resultando em um número maior de regiões regimentadas. A Figura 11 apresenta a segmentação da imagem de um nariz de bolha alongada utilizando o algoritmo K-Means com diferentes valores de k . As imagens segmentadas apresentam a divisão da imagem original em grupos de intensidade de pixels para $k = 2$, $k = 4$ e $k = 6$.

Figura 11 - Segmentação da imagem utilizando o algoritmo K-means



Fonte: Autoria própria.

Os primeiros métodos de segmentação por agrupamento surgiram no final da década de 1970. Um dos trabalhos iniciais foi o de Schachter e Dams (1978) que aplicou *feature clustering* em imagens em tons de cinza, destacando a importância da escolha dos atributos (como intensidade e coordenadas espaciais) mais do que do algoritmo em si. Em seguida, Silverman e Cooper (1988) propuseram um algoritmo aglomerativo que utilizava a distância de Mahalanobis para fundir clusters de coeficientes de modelos polinomiais da imagem, o que permitia segmentar regiões a partir de representações matemáticas locais. Já Wu e Leahy (1993) introduziram uma abordagem baseada em teoria de grafos, tratando os pixels como nós conectados por medidas de similaridade e identificando segmentos pela remoção de arestas, o que aproximou a segmentação de uma lógica de detecção de contornos.

Esses trabalhos estabeleceram as bases para o desenvolvimento de métodos posteriores, cada vez mais sofisticados, incluindo o uso de redes neurais e modelos probabilísticos. Nas últimas décadas, a segmentação por agrupamento evoluiu de

métodos clássicos para modelos baseados em redes neurais que integram aprendizado de representação e clusterização. Um exemplo recente de aplicação de *clustering* em segmentação de imagens é o *Invariant Information Clustering* (IIC), que aprende a separar automaticamente regiões de interesse usando apenas dados não rotulados, evitando problemas comuns de outros métodos e alcançando resultados de ponta em benchmarks de segmentação Ji et al. (2019).

2.2.2.6 Técnicas baseadas em aprendizado de máquina

As técnicas clássicas de segmentação de imagens, como limiarização, detecção de bordas, crescimento de regiões e transformada Watershed, foram amplamente utilizadas por sua simplicidade e eficiência computacional Sharaf (2024). No entanto, esses métodos dependem de ajuste manual de parâmetros e apresentam desempenho limitado frente a variações nas características das imagens. Para superar tais dificuldades, técnicas baseadas em aprendizado de máquina têm sido amplamente exploradas, com propósito de tornar a segmentação de imagem mais autônoma, adaptativa e robusta. Nessa abordagem, em vez de definir manualmente critérios e limiares, o modelo aprende os padrões e relações automaticamente a partir de exemplos. Essa abordagem se mostrou extremamente eficaz em tarefas de visão computacional, como reconhecimento de objetos, detecção de bordas e segmentação (Lecun et al., 2015).

As técnicas de aprendizado de máquina representam uma mudança de paradigma, de sistemas baseados em regras explícitas para sistemas capazes de aprender padrões através de dados. Segundo Goodfellow et al. (2015), esses métodos podem ser organizados em três principais paradigmas de aprendizado:

- Não supervisionado: os modelos processam conjuntos de dados com diversas características e buscam identificar padrões e regularidades não aparentes que descrevem a organização interna dos dados. Durante o treinamento, os parâmetros são ajustados para minimizar uma função de custo interna, sem utilizar rótulos ou saídas de referência.
- Supervisionado: os modelos também processam conjuntos de dados com diversas características, mas cada exemplo usado para treinamento do modelo está associado a um rótulo. Durante o treinamento, os parâmetros

atualizados com base nos pares de dados de entrada e saída, minimizando uma função de perda que quantifica a diferença entre a predição do modelo e o valor de referência.

- Por reforço: nessa abordagem os modelos aprendem com interação contínua com o ambiente, sem orientação direta, realizando ações e observando as suas consequências. Durante o treinamento, os agentes ajustam seus parâmetros com base em recompensas ou punições relacionadas aos resultados de suas ações no ambiente, buscando maximizar a recompensa acumulada.

Entre esses paradigmas, se destaca o aprendizado supervisionado, consolidado como base dos modelos modernos de segmentação, como o modelo U-Net ou Segment Anything Model (SAM), utilizados neste trabalho. As abordagens modernas de aprendizado de máquina têm origens nos fundamentos teóricos de redes neurais. O primeiro modelo matemático de um neurônio artificial foi desenvolvido por Mcculloch e Pitts (1943), dando início à formulação de sistemas computacionais inspirados no funcionamento biológico do cérebro. Nesse modelo pioneiro, os neurônios foram representados como unidades binárias capazes de somar entradas discretas e aplicar uma função limiar, permitindo, em teoria, que qualquer função lógica fosse representada.

Alguns anos depois, Hebb (1949) propôs um modelo de aprendizado inspirado em mecanismos biológicos, conhecida como regra de aprendizado Hebbiana. Esse conceito introduziu o aprendizado baseado em experiência, baseando-se no princípio de que neurônios que disparam juntos reforçam suas ligações. Dando continuidade ao avanço das redes neurais, Rosenblatt (1958) desenvolveu o Perceptron, o primeiro modelo computacional capaz de realizar aprendizado supervisionado. Esse modelo ajustava automaticamente os pesos das conexões com base no erro entre a saída prevista e o resultado desejado, permitindo que o sistema aprendesse a classificar padrões a partir de exemplos.

Apesar do avanço representado pelo Perceptron, Minsky e Papert (1969) identificaram limitações teóricas significativas nesse modelo. Em especial, demonstraram que o Perceptron não era capaz de resolver problemas não lineares, uma vez que sua arquitetura de camada única só permitia a geração de fronteiras

lineares de decisão. Essas restrições levaram a um período de ceticismo da comunidade científica e redução no financiamento de pesquisas de redes neurais.

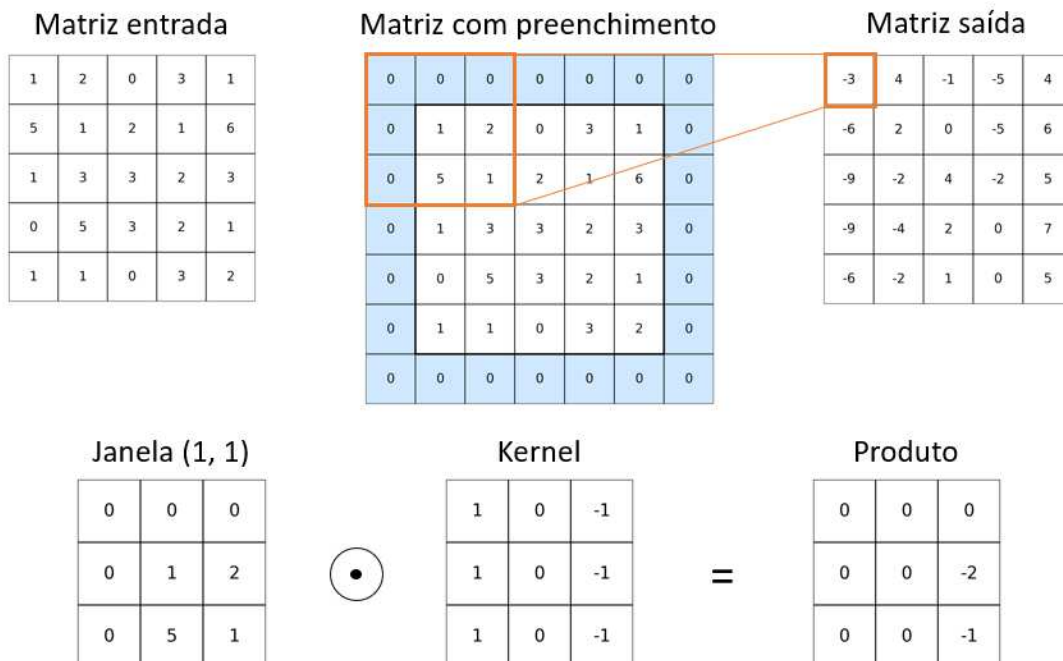
Nas décadas seguintes, com maior poder computacional e novas abordagens de aprendizado supervisionado, o campo voltou a crescer de forma expressiva. Nesse contexto, consolidou-se o conceito de Redes Neurais Artificiais (*Artificial Neural Networks* - ANNs) como modelos computacionais compostos por unidades interconectadas (neurônios artificiais), capazes de aprender padrões complexos a partir de dados de entrada (O'Shea; Nash, 2015). Entre essas arquiteturas, destacam-se as Redes Neurais Multicamadas (*Multilayer Perceptrons* - MLPs), que estendem o Perceptron original ao introduzir camadas ocultas capazes de representar relações não lineares. Essa estrutura foi consolidada com o trabalho de Rumelhart, Hinton e Williams (1986), que introduziram o algoritmo de retropropagação de erro (*backpropagation*), possibilitando o treinamento eficiente de redes com múltiplas camadas. Entretanto, as MLPs possuem uma limitação fundamental quando aplicadas a dados com estrutura espacial ou temporal, como imagens ou séries temporais, pois tratam cada elemento de forma independente, desconsiderando as relações locais e temporais entre seus elementos próximos. Isso faz com que o número de parâmetros cresça exponencialmente com o número de elementos do dado, exigindo um poder computacional maior e levando a perda de informação sobre padrões locais.

Para superar tais limitações, as Redes Neurais Convolucionais (*Convolutional Neural Networks* - CNNs) foram introduzidas por Lecun et al. (1989), capazes de preservar a estrutura espacial dos dados e extrair características locais. Essas redes são organizadas em camadas (*layers*), que podem ser definidas como os níveis de transformações sucessivas dos dados de entrada. Cada camada recebe uma entrada, aplica uma operação matemática e produz uma saída que serve de entrada para a próxima camada. Essas camadas aplicam operações específicas, como convoluções, ativações e reduções, construindo uma representação hierárquica da informação. As primeiras camadas capturam padrões mais simples, enquanto as camadas mais profundas combinam esses elementos em estruturas mais complexas. Diferentemente das MLPs, as CNNs incorporam as camadas convolucionais, que aplicam filtros (*kernels*) sobre múltiplos elementos dos dados de entrada, explorando as correlações espaciais dos dados de entrada.

Uma camada convolucional aprende filtros (*kernels*) que extraem padrões locais dos dados. A operação convolucional consiste na translação do kernel sobre o

domínio da entrada, com preenchimento (*padding*) e passo (*stride*) especificados. O tamanho espacial do kernel é um parâmetro do modelo e sua profundidade, na maioria das aplicações, coincide com o número de canais da entrada. Devido ao compartilhamento de pesos, cada filtro aprendido é aplicado para todas as posições da entrada, gerando um mapa de características (também chamado de mapa de ativação ou canal de saída). Em uma mesma camada convolucional, diferentes filtros podem ser aprendidos, gerando diferentes mapas de características para mesma entrada, e esses mapas são empilhados para formar o tensor de saída. É comum empregar camadas de normalização por lote (*batch normalization*) e funções de ativação não lineares, que tem como objetivo, respectivamente, estabilizar a escadas ativações por canal e acrescentar a não linearidade aos modelos. A Figura 12 ilustra a operação de convolução com preenchimento utilizando um kernel de 3x3. A entrada possui um único canal e gera um mapa de saída de mesmo tamanho. Um preenchimento é adicionado à matriz de entrada, e cada janela é multiplicada elemento a elemento pelo kernel 3x3. Em seguida, todos os valores resultados dessa multiplicação são somados para compor o valor correspondente no mapa de saída.

Figura 12 – Exemplo de convolução



$$\text{soma (Janela} \odot \text{Kernel)} = 0 + 0 + 0 + 0 + 0 - 2 + 0 + 0 - 1 = -3$$

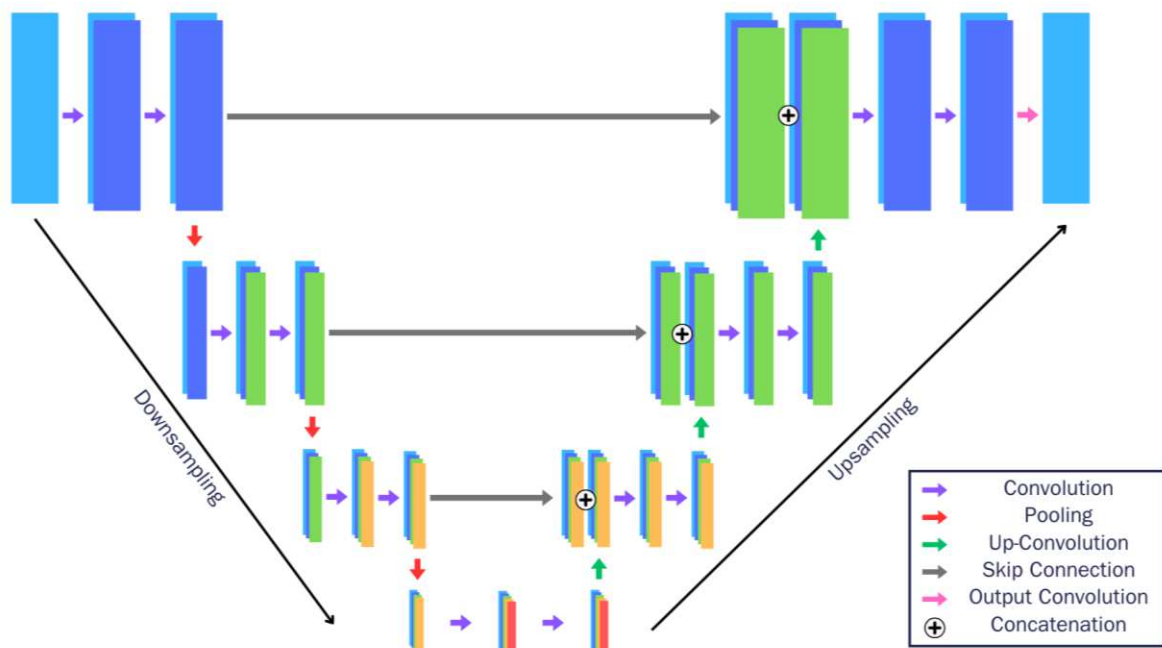
Fonte: Autoria própria.

As redes neurais convolucionais são amplamente utilizadas em tarefas de visão computacional e servem de base para arquiteturas mais avançadas, como a U-Net, proposta por Ronneberger, Fischer e Brox (2015), desenvolvida originalmente para segmentação semântica de imagens biomédicas. O modelo U-Net expande a estruturas das redes neurais tradicionais, ao adotar uma arquitetura simétrica do tipo *encoder-decoder*. Nesse tipo de arquitetura, o *encoder* realiza a extração hierárquica de características e o *decoder* reconstrói a imagem segmentada, recuperando a resolução original da imagem de entrada. Para preservar as informações espaciais perdidas durante o processo de geração do mapa de características, as conexões de salto (*skip connections*) interligam cada camada do *encoder* as camadas do *decoder*, mantendo uma segmentação precisa e detalhada. Esse tipo de arquitetura é amplamente utilizado em diferentes áreas de aprendizado de máquina, abrangendo desde modelos de processamento de linguagem natural, como o modelo de tradução de sentenças apresentados por Sutskever, Vinyals e Le (2014), até modelos de processamento de áudio e fala como o Tacotron apresentado por Wang et al. (2017). Essa diversidade de aplicações evidencia a versatilidade do paradigma *encoder-decoder*, capaz de generalizar domínios de natureza diversa, como visão computacional, linguagem e áudio.

A arquitetura U-Net é apresentada a seguir, conforme o diagrama da Figura 13. As etapas do fluxo de dados são detalhadas, destacando-se a função de cada bloco convolucional e suas conexões. Esse tipo de arquitetura organizada em três partes principais: *encoder (downsampling)*, *bottleneck* e *decoder (upsampling)*. O *encoder* recebe a imagem de entrada e reduz progressivamente a resolução espacial e aumenta o número de mapas de características. Em todos os níveis de resolução do *encoder*, são aplicadas convoluções, um bloco função de ativação e uma camada de *pooling* no final. Na maioria das aplicações, a resolução diminui pela metade e o número de filtros é duplicado em relação a camada anterior. O *bottleneck* atua como a conexão entre o *encoder* e o *decoder*, sendo composto por camadas convolucionais com o maior número de filtros e menor resolução, representando as características de alto nível da imagem de entrada. O *decoder* executa o caminho inverso ao do *downsampling*, reconstruindo gradualmente a resolução espacial enquanto a profundidade dos mapas de características, até gerar a imagem segmentada. Inicialmente, em cada nível de resolução do *upsampling*, a resolução espacial é dobrada em relação ao nível anterior e, esse resultado é concatenado com o mapa de

características do nível correspondente do *encoder*, com objetivo de recuperar detalhes perdidos durante o *pooling*. Após a concatenação, duas convoluções são aplicadas, gerando o mapa de características daquele nível de resolução. A etapa final do *decoder* consiste em uma convolução 1x1, responsável por transformar os mapas de características do último nível de resolução em uma imagem segmentada com o mesmo tamanho da imagem de entrada.

Figura 13 – Diagrama arquitetura U-Net



Fonte: Autoria própria.

Desde sua introdução, a U-Net vem sendo empregada como base para novos modelos, mantendo a base *encoder-decoder* e incorporando mecanismos modernos de redes neurais, como atenção, blocos residuais e Transformers. Um exemplo é a U-Net ++, proposta por Zhou et al. (2018), que introduz convoluções intermediárias nas *skip connections* com intuito de reduzir o descompasso semântico entre o *encoder* e o *decoder*, resultando em melhor desempenho nos critérios de avaliação. Outro exemplo relevante é a Attention U-Net (Oktay et al., 2018), que incorpora módulos de atenção (*attention gates*) nas *skip connections*, com objetivo de selecionar as características mais relevantes durante a reconstrução da imagem segmentada. Outros modelos, como Milletari, Navab e Ahmadi (2016) e Çiçek et al. (2016), estendem a arquitetura original para dados volumétricos (3D), permitindo sua

aplicação em imagens médicas, como tomografias e exames de ressonância magnética.

A etapa de avaliação é essencial para medir o desempenho de modelos de segmentação baseados em aprendizado de máquina. Geralmente, em tarefas de classificação, a acurácia é a métrica utilizada para indicar o número de acertos globais. Normalmente, no caso de segmentação de imagens, há um desbalanceamento no número de pixels pertencentes ao fundo (*background*) e os objetos de interesse. Por essa razão, a acurácia pode superestimar o desempenho do modelo, pois leva em conta o acerto dos pixels do fundo. Logo, para segmentação de imagens, as métricas mais utilizadas são baseadas em sobreposição e equilíbrio de classes. De acordo com Müller, Soto-Rey e Kramer (2022), as métricas de acurácia, coeficiente de Dice e interseção sobre união (*Intersection over Union* - IoU) são definidas com base nos valores de verdadeiros positivos (VP), falsos positivos (FP), verdadeiros negativos (VN) e falsos negativos (FN), conforme as Equações 14 a 16.

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FN + FP} \quad (14)$$

$$\text{DICE} = \frac{2TP}{2TP + FP + FN} \quad (151)$$

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (16)$$

Embora arquiteturas como a U-Net e suas variações tenham consolidado o uso de redes neurais para segmentação de imagens, os avanços recentes na área de processamento de imagem caminham em direção a modelos mais generalistas e escaláveis. Conforme Bommasani et al. (2022), os chamados *foundation models* (modelos fundacionais) distinguem-se por serem treinados em grandes volumes de dados e por apresentarem ampla capacidade de adaptação, podendo ser ajustados para diferentes tarefas sem a necessidade de um novo treinamento completo. No contexto de visão computacional, exemplos como o Segment Anything Model (Kirillov et al., 2023), o CLIP (Radford et al., 2021) e o SEEM (Zou et al., 2023), demonstram potencial para segmentação e reconhecimento de objetos em diferentes domínios, com baixa supervisão e alta generalização. O Segment Anything Model (SAM) foi treinado com o conjunto de dados de segmentação SA-1B, criado pelos próprios autores do trabalho. Esse conjunto contém 11 milhões de imagens e 1,1 bilhão de

máscaras anotadas, abrangendo uma ampla variedade de domínios visuais. Essas características, tornam o SAM um modelo que pode ser usado como uma ferramenta eficaz de anotação automática, capaz de gerar máscaras de treinamento para outros modelos de segmentação, reduzindo significativamente o custo e o tempo associado a geração de máscaras de treinamento.

Os modelos de segmentação baseados em redes neurais têm se destacado por sua capacidade em diferentes contextos de imagem, contudo, esses modelos exigem grandes volumes de dados anotados para alcançar bom desempenho. A anotação manual de máscaras constitui uma etapa complexa e sujeita a inconsistências, pois requer a delimitação detalhada pixel a pixel. Um exemplo que evidencia a necessidade de máscaras anotadas em grande quantidade para o treinamento de modelos de segmentação é o de Benameur et al. (2024). O estudo empregou 8.400 imagens divididas em 70% para treino, 10% validação e 20% teste, alcançando Dice de 0,965 e 0,945, nas fases diastólica e sistólica (diferentes fases do ciclo cardíaco), respectivamente. O próprio trabalho de Ronneberger, Fischer e Brox (2015) evidencia que, embora a U-Net obtenha bom desempenho com poucas imagens quando associada a *data augment*, a qualidade e quantidade de máscaras anotadas são fatores determinantes para o sucesso do modelo de segmentação.

2.3 Rastreamento de objetos

O rastreamento de objetos é uma etapa central em tarefas de visão computacional. Ele pode ser definido como o processo de detectar e rastrear objetos de interesse em uma sequência de imagens, com objetivo de estimar características como trajetória, posição e estado dos objetos (Yilmaz; Javed; Shah, 2006). O rastreamento de objetos se mostra essencial para análise de imagens em aplicações como análise do movimento de bolhas (Q. Wang et al., 2021), sistemas de monitoramento de segurança (Mathur; Somwanshi; Bundele, 2018), e rastreamento de transporte e tráfego, veículos autônomos (Hoffmann et al., 2021), entre outras. O rastreamento de objetos está sujeito a diversos empecilhos, devido a presença de ruídos, artefatos visuais, oclusões, fragmentação ou aglomeração incorreta de objetos, além de mudanças abruptas de movimento ou forma dos objetos. Em aplicações envolvendo escoamento bifásico, esses problemas podem ser agravados devido ao grande número de bolhas presentes em cada frame, pela coalescência e

desprendimento de bolhas, além do fato de que, muitas vezes, as bolhas possuem dimensões superiores ao campo de visão das imagens adquiridas.

Para lidar com essa grande diversidade de cenários e condições presentes nas tarefas de rastreamento de objetos de rastreamento de objetos, diferentes metodologias foram desenvolvidas. De acordo com Yilmaz, Javed e Shah (2006), as técnicas de rastreamento de objetos podem ser distintas em três categorias principais, definidas de acordo com a forma como o objeto é representado:

- Baseada em ponto (*point tracking*): Os objetos detectados são representados por pontos característicos, geralmente esses pontos correspondem a posição central (centroide) ou marcadores específicos. Essa metodologia depende da detecção externa dos objetos, e o rastreamento é obtido através da associação desses pontos ao longo do tempo, com base no estado anterior do objeto (por exemplo, posição e velocidade).
- Baseada em kernel (*kernel tracking*): os objetos são representados por um modelo de forma e aparência, denominado kernel. Esse kernel pode assumir diferentes formatos e está associado a uma descrição interna de suas características. O rastreamento de objetos é realizado com base em medidas de similaridade e transformações paramétricas aplicadas entre os kernels em frames consecutivos.
- Baseado em silhuetas (*silhouette tracking*): os objetos são representados pela silhueta completa, ou seja, pela distribuição espacial dos pixels pertencentes aos objetos. O rastreamento de objetos é realizado pela estimativa da região e da forma do objeto ao longo dos frames consecutivos (Luo et al., 2022).

Além da classificação baseada na representação dos objetos, os métodos de rastreamento de objetos também podem ser diferenciados segundo a estratégia de associação entre os objetos ao longo do tempo. Essa abordagem pode envolver modelos de predição, regras geométricas, otimização global ou modelos de aprendizado. Neste contexto, se destaca o rastreamento baseado em detecção (*Detection-Based Tracking* - DBT), em que consiste em identificar os objetos por frame e associar correspondências temporais entre essas detecções (Luo et al., 2022). Esse paradigma, também denominado *tracking-by-detection*, é amplamente aplicado em

tarefas de rastreamento de múltiplos objetos (*Multiple Object Tracking - MOT*), cujo objetivo é identificar e rastrear, simultaneamente, diversos objetos ao longo de uma sequência de frames. O principal desafio consiste em associar as detecções correspondentes a múltiplos objetos, mesmo em cenários com oclusões, interações e detecções incorretas (Luo et al., 2022). Essa associação é geralmente baseada em uma combinação entre modelos probabilísticos de movimento, como o Filtro de Kalman (Kalman, 1960), e algoritmos de otimização global como o algoritmo de Hungarian (Kuhn, 1955). Essa metodologia é empregada em rastreadores baseados em detecção, como o SORT (Bewley et al., 2017) e o DeepSORT (Wojke et al., 2017), que são amplamente adotados por alinharem simplicidade, eficiência computacional e boa precisão temporal.

Em rastreadores baseados em detecção, o desempenho da associação temporal está relacionado a capacidade do sistema prever o movimento dos objetos, mesmo na presença de ruído e perda temporária de detecção. Para isso, os filtros probabilísticos são empregados, que são capazes de estimar o estado de um sistema dinâmico de forma recursiva, combinando etapas de predição e correção para lidar com ruídos e incertezas. Na etapa de predição, com base no estado anterior, o modelo de movimento é utilizado para estimar o estado futuro do sistema e sua incerteza. O vetor de estados representado por um conjunto de variáveis que descrevem o estado atual do objeto rastreado, geralmente composto por suas coordenadas espaciais, dimensões e componente de velocidade ou aceleração. O filtro de Kalman é uma simplificação dessa estrutura, fornecendo uma solução ótima sob a suposição de que o modelo é linear e o ruído possui distribuição Gaussiana. O modelo de movimento do Filtro de Kalman pode ser definido pelas equações:

$$x_k = Fx_{k-1} + Gw_{k-1} \quad (172)$$

$$P_k = FP_{k-1}F^T + GQG^T \quad (18)$$

Onde:

- x : vetor de estado;
- w : ruído do processo;
- F : matriz de transição de estado;
- G : matriz de entrada do ruído de processo;

- Q : matriz de covariância do ruído de processo;
- P : matriz de covariância do erro de estimativa;
- k : instante de tempo atual;
- $k - 1$: instante de tempo anterior.

Em seguida, na etapa de correção, a estimativa é atualizada com base na nova medição, ponderando a confiabilidade relativa entre o valor estimado na etapa de predição e a medição observada. O grau dessa ponderação é definido matematicamente pelo ganho de Kalman (K_k), obtido a partir das incertezas associadas ao modelo de predição e o sistema de observação.

A atualização do estado é expressa pela Equação 19:

$$\hat{x}_k = \hat{x}_{\bar{k}} + K_k(z_k - H\hat{x}_{\bar{k}}) \quad (19)$$

Onde:

- \hat{x}_k : vetor de estado corrigido;
- $\hat{x}_{\bar{k}}$: vetor de estado predito;
- z_k : medição observada;
- H : matriz de transição de estado;
- K_k : ganho de Kalman.

Uma vez calculado o ganho, o filtro atualiza o estado, combinando a predição e a medição de forma ponderada. O resultado é uma nova estimativa, baseada na predição e na observação. Por fim, ocorre a atualização da incerteza, assim o filtro executa iterativamente as etapas de predição, correção a atualização, refinando as estimativas do estado ao longo do tempo. No contexto do

No contexto do SORT, as medições utilizadas na etapa de correção correspondem as detecções obtidas no frame atual, as quais são associadas às predições do filtro de Kalman através do algoritmo Húngaro. Para isso, é construída uma matriz de custo C_{ij} , onde cada elemento representa o custo de associar a predição i à detecção j . A função de custo define o valor associado a cada possível correspondência entre predições e detecções, podendo considerar diferentes critérios de similaridade, como distâncias espaciais entre centroides, grau de sobreposição

entre as caixas delimitadoras ou velocidade dos objetos detectados. No caso específico do SORT, a função de custo é definida a partir da Intersecção sobre União (IoU), entre as caixas delimitadoras preditas e detectadas. Além disso, é adotado um limiar mínimo de IoU para rejeitar associações com baixa sobreposição e evitar correspondências incorretas. Após o cálculo da função de custo, o Algoritmo Húngaro percorre todas as combinações possíveis de associações e seleciona aquela que resulta no menor somatório global de custos, garantindo que cada predição seja associada a, no máximo, uma única detecção.

2.4 Trabalhos relacionados

Um dos primeiros estudos a empregar imagens de escoamentos bifásicos gás-líquido no padrão intermitente foi desenvolvido por Davies e Taylor (1949), no qual foi aplicada uma técnica manual para medir e modelar a geometria e velocidade de subida de bolhas de ar em líquidos. Posteriormente, com o avanço das técnicas computacionais, Schrodtt e Saunderst (1981) desenvolveram um sistema de processamento de imagens baseado em detecção de bordas para segmentar as regiões correspondentes a fases líquida e gasosa, possibilitando a extração parâmetros geométricos, tais como área, comprimento e diâmetro de bolhas. Os trabalhos de Gopal (1997), Dinht, Kim e Choi (1999) e Polonsky, Barnea e Shemer (1999), são exemplos das primeiras aplicações estruturadas de processamento de imagens ao estudo do escoamento bifásico em regime de golfadas para tubulações horizontais e verticais. O trabalho de Gopal (1997) desenvolveu uma metodologia baseada em detecção manual de bordas, utilizando uma bancada experimental com uma tubulação horizontal com diâmetro de 75 mm e escoamentos água-CO₂ e óleo e CO₂, sendo um dos primeiros trabalhos a obter perfis locais de velocidade e fração de vazio, além de propor um modelo dinâmico para *holdup* de líquido. Dinht (1999) e Polonsky, Barnea e Shemer (1999) aplicaram técnicas de processamento de imagens baseadas em filtragem, limiarização e detecção de bordas em tubulações verticais em escoamentos água-ar, permitindo a segmentação automáticas de bolhas e a extração de parâmetros geométricos e dinâmicos relevantes.

Dando continuidade a essa linha de investigação, Mayor, Pinto e Campos (2007) propuseram uma metodologia para processamento de imagens de escoamento bifásico em regime de golfadas estruturada e reproduzível, baseada em técnicas como subtração de fundo, conversão para escala de cinza, limiarização binária, filtragem

mediana e erosão morfológicas. Além disso, o trabalho também se destacou pela condução de uma análise de incertezas no processamento de imagens, evidenciando que era possível atingir erros relativos inferiores a 10% para estimativa dos principais parâmetros do escoamento bifásico em regime de golfadas. Outros trabalhos, como os de Oliveira et al. (2015) e Widyatama et al. (2018), desenvolveram metodologias para escoamento horizontal bifásico gás-líquido em regime de golfadas, empregando técnicas de segmentação semelhantes às de Mayor. Oliveira et al. (2015) avançaram ao introduzir a caracterização geométrica detalhada do nariz e da cauda de bolhas alongadas, indicando uma relação linear entre a posição e formato do nariz com o número de Froude. Por sua vez, a metodologia desenvolvida por Widyatama et al. (2018), baseada em filtros *top-hat* e *bottom-hat*, além da subtração de fundo, normalização de intensidade, binarização e extração de contornos. Um aspecto relevante da metodologia consiste na remoção de bolhas dispersas presentes na região de cauda das bolhas alongadas, possibilitando obter uma definição mais robusta do contorno da cauda.

Além dessas abordagens clássicas baseadas em operações morfológicas, outros trabalhos passaram a explorar metodologias distintas para segmentação de imagens de escoamentos bifásicos horizontais em regime de golfadas. Dentre eles destacam-se os estudos de Coelho (2011), do Amaral et al. (2013) e Mohammed et al. (2016). Coelho (2011) introduziu o uso de método Leve-Set para segmentação de imagens, integrando a extração de parâmetros e classificação de padrões com base nessas imagens segmentadas. Em uma abordagem distinta, Mohammed et al. (2016) empregaram o uso de *Approximate Median Method* (AMM) e operações morfológicas para segmentação, permitindo a separação eficiente entre a região de fundo estático e as bolhas e interfaces gás-líquido. Por sua vez, do Amaral et al. (2013) propuseram uma cadeia de segmentação operadores morfológicos avançados (*top-hat* e *bottom-hat*) e a transformada de Watershed. Ressalta-se ainda que o trabalho apresenta uma metodologia de agregação temporal de imagens para reconstrução de bolhas alongadas maiores que a janela de visualização, que serviu de base conceitual para metodologia proposta neste trabalho.

Embora métodos baseados em filtros, limiarização, operações morfológicas e transformadas tenham sido amplamente utilizadas, apresentando bons resultados em condições controladas, seu desempenho pode ser limitado em situações de ruído elevado, interfaces complexas e ampla variação das condições de escoamento. A

necessidade de calibração específica para cada condição pode limitar a aplicabilidade em estudos sistemáticos e automatizados. Essas dificuldades impulsionaram a adoção de abordagens baseadas em aprendizado profundo, que permitem capturar padrões complexos através de treinamento. A partir desse cenário, redes neurais convolucionais, por exemplo a U-Net (Zhou et al., 2018), passaram a ser utilizadas para segmentação de imagens em escoamentos bifásicos. Alguns trabalhos recentes que utilizam essas abordagens desenvolvidos por Olbrich et al. (2022), Da Silva et al. (2024) e Pinto et al. (2023). Os autores concluem que modelos baseados em arquitetura U-Net podem segmentar de forma automática as regiões de gás e líquido, mesmo para cenários complexos, alcançando boas métricas de desempenho. A Tabela 1 apresenta uma comparação entre esses trabalhos, os quais empregam arquitetura U-Net e utilizam imagens anotadas manualmente para o treinamento dos modelos. Além das arquiteturas do tipo U-Net, estudos mais recentes também têm explorado outras abordagens para detecção e rastreamento de bolhas. Nesse contexto, Neves et al. (2024) aplicaram o modelo fundacional *Segment Anything Model* (SAM), para segmentação automática de bolhas dispersas sem necessidade de treinamento, enquanto, Metzker et al. (2025) utilizaram o modelo YOLO v8 (Yaseen, 2024) em conjunto com técnicas de *optical flow*, para detecção e rastreamento de bolhas de Taylor. Em resumo, todos esses trabalhos evidenciam uma tendência do uso de técnicas baseadas em aprendizado profundo para segmentação, detecção e caracterização de escoamentos bifásicos. Entretanto, é importante ressaltar os desafios relacionados à necessidade de grandes conjuntos de dados para o treinamento desses modelos e ao elevado custo computacional associado.

Tabela 1 - Comparação entre estudos que aplicaram U-Net à segmentação em escoamentos bifásicos

Artigo	Dataset	Treinamento	Teste	Métrica	Resultado
Oldrich et al. (2022)	544 pares imagem-máscara	88,88%	11,20%	Acurácia	97,91% (treino) / 97,74% (teste)
Ahmad et al. (2022)	648 pares imagem-máscara	100%	Validação qualitativa	Acurácia	95.69% (treino)
Pinto et al. (2023)	1161 pares imagem-máscara	60% - 80%	20% - 40%	Dice	86,79%–99,81% (teste)
da Silva et al. (2024)	600 pares de sub-imagens (128×128 px), gerados a partir de 6 pares imagem-máscara	70%	30%	Acurácia	71.3% (treino) / 70.0% (teste)

Fonte: Autoria própria.

3 MATERIAIS E MÉTODOS

Esta seção descreve a metodologia empregada para análise de escoamento bifásico horizontal no padrão intermitente, abrangendo desde a aquisição de imagens até o processamento para extração dos parâmetros característicos do escoamento. O procedimento proposto combina diferentes técnicas de visão computacional e aprendizado profundo para segmentar, identificar e rastrear as bolhas que compõem o escoamento bifásico gás-líquido.

Os dados experimentais utilizados neste trabalho foram obtidos a partir de quatro baterias experimentais. Inicialmente, uma bateria preliminar foi executada a 3 bar, correspondendo ao primeiro teste da bancada experimental. Essa bateria foi conduzida com objetivo de explorar os limites operacionais da bancada em termos de velocidades superficiais do líquido (J_L) e do gás (J_G), bem como obter uma avaliação inicial das condições em que o escoamento poderia se estabelecer no padrão de golfadas. As imagens foram adquiridas a uma resolução de 720x140 pixels a uma frequência de 250 Hz. Para essa bateria, foi realizado um processamento manual das imagens com objetivo de gerar valores de referência, como a velocidade e a frequência de passagem de bolhas alongadas, utilizadas na validação inicial da metodologia de segmentação. Para cada ponto experimental, as bolhas alongadas foram identificadas visualmente, possibilitando o cálculo da frequência de ocorrência. A velocidade dessas bolhas foi determinada a partir da marcação, por cliques do operador, do primeiro e do último frame em que o nariz da bolha aparece no campo de visão, sendo determinada como a razão entre o deslocamento espacial do nariz e o intervalo de tempo entre os dois frames. A Tabela 2 apresenta as diferentes combinações de velocidades superficiais de gás (J_G) e de líquido (J_L) para bateria a 3 bar. Na tabela também são apresentados o número de bolhas alongadas identificadas e o respectivo tempo de aquisição de imagens do ponto experimental.

Tabela 2 – Bateria experimental a 3 bar

Ponto	J_G [m/s]	J_L [m/s]	Número de bolhas	Tempo [s]
P01	0,02	0,10	29	58,3
P03	0,22	0,10	21	56,1
P04	0,16	0,11	15	51,0
P05	0,36	0,12	9	40,4
P09	0,10	0,21	43	54,6
P10	0,19	0,20	43	57,9

P11	0,28	0,20	38	54,6
P12	0,37	0,20	25	39,8
P13	0,45	0,20	8	20,8
P17	0,11	0,31	74	55,5
P18	0,19	0,31	70	59,4
P19	0,29	0,30	54	52,0
P20	0,37	0,30	32	39,6
P21	0,48	0,30	22	30,8
P22	0,52	0,30	9	19,4
P25	0,12	0,44	77	56,0
P26	0,22	0,43	87	58,4
P27	0,26	0,44	46	38,6
P28	0,37	0,43	45	43,9
P29	0,51	0,42	28	35,0
P30	0,58	0,46	16	25,2
P33	0,12	0,52	103	66,2
P34	0,22	0,52	98	59,9
P35	0,30	0,52	81	57,6
P36	0,36	0,52	52	43,7
P37	0,49	0,52	35	37,4
P38	0,59	0,52	21	29,1
P60	0,03	0,12	22	48,0
P61	0,08	0,11	33	49,7
P62	0,03	0,21	30	49,1
P63	0,13	0,15	41	41,3
P64	0,03	0,31	46	53,4
P65	0,07	0,31	61	41,0
P66	0,05	0,40	54	41,5
P67	0,06	0,40	70	43,1
P68	0,04	0,50	57	43,3
P69	0,07	0,50	43	43,4
P70	0,06	0,06	11	39,1
P71	0,08	0,07	17	40,8

Fonte: Autoria própria.

Após essa etapa inicial, foram adquiridas as demais baterias experimentais à pressão de 4 bar. Cada bateria de 4 bar contempla diferentes combinações de velocidades superficiais de gás (J_G) e de líquido (J_L), abrangendo uma ampla faixa de padrões de escoamento intermitente, apresentados na Tabela 5 do Anexo A. Essas combinações foram definidas com base na bateria exploratória de forma a representar tanto condições típicas de escoamento em golfadas quanto situações próximas às regiões de transição entre padrões. Cada ponto experimental corresponde a 4 minutos

de gravação. As imagens foram adquiridas a 500 Hz e com uma resolução de 1440x360 pixels.

A metodologia está organizada em quatro etapas. Primeiramente, o circuito experimental é descrito, incluindo as dimensões da bancada, fluidos utilizados e o sistema de aquisição de imagens de alta velocidade. Em seguida, a etapa de segmentação de imagens é apresentada, destinada a distinguir as fases líquida e gasosa. Nessa etapa, a arquitetura do modelo e o conjunto de dados utilizado são apresentados. A arquitetura adotada é uma rede neural do tipo U-Net, que foi treinada com máscaras geradas automaticamente pelo modelo Segment Anything Model (Ravi et al., 2024).

A terceira etapa está relacionada ao rastreamento de objetos, com objetivo de associar as múltiplas bolhas segmentadas ao longo dos frames. Para isso, empregou-se o algoritmo Húngaro, associado a diferentes critérios de associação (*gates*) e distância de Mahalanobis, juntamente com um filtro de Kalman destinado a suavização de trajetórias e à recuperação de identificações perdidas.

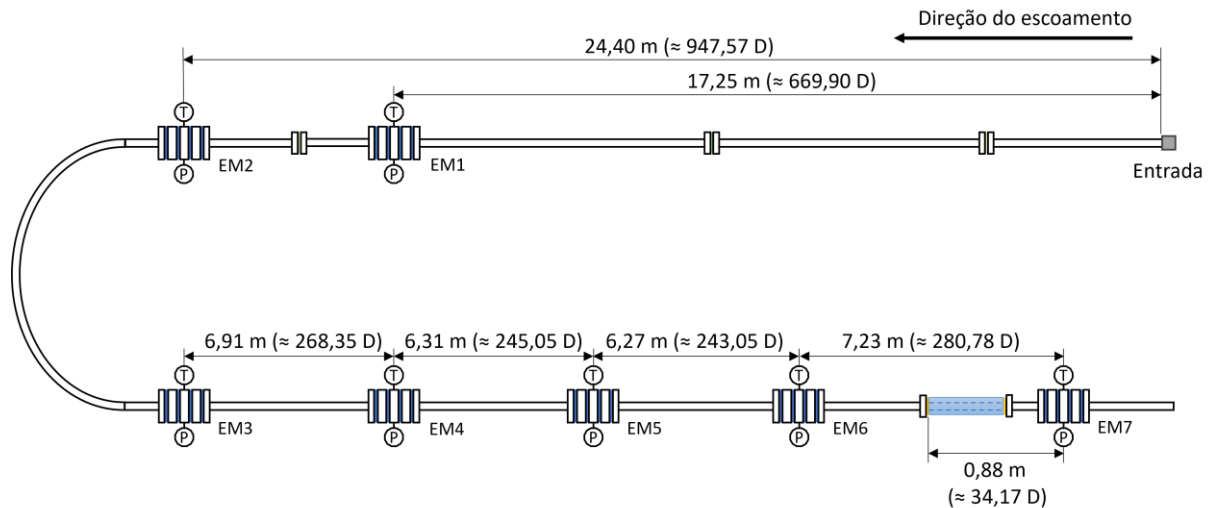
Por fim, na última etapa, as bolhas alongadas são filtradas e agregadas para identificação das células unitárias, possibilitando a extração dos parâmetros do escoamento, incluindo velocidade de bolha, frequência, altura de nariz de bolha, fração de vazio e comprimentos (célula unitária, bolha e pistão de líquido).

3.1 Circuito experimental

O circuito experimental utilizado foi originalmente projetado por Gmyterco (2018) para emular o comportamento de escoamentos bifásicos gás-líquido em condições de alta pressão e diferentes velocidades superficiais. Um diagrama do circuito é apresentado na Figura 14. O sistema é composto por três seções principais: a seção de líquido e a seção de testes. A fase líquida é pressurizada por meio de uma bomba, enquanto a fase gasosa é pressurizada por um compressor. Ambas as seções dispõem de reservatórios, trocadores de calor e medidores de vazão do tipo Coriolis, permitindo o controle da temperatura em ambas as linhas, bem como a medição da vazão de cada fase antes da entrada no misturador. Para emular o comportamento dos cenários do pré-sal, o gás hexafluoreto de enxofre (SF_6) foi escolhido devido sua densidade elevada. Dessa forma, pode-se alcançar uma razão entre as fases semelhante à de sistemas reais, mesmo operando em pressões mais baixas. A seção de testes é composta por uma tubulação horizontal de aço com diâmetro interno de

26 mm e 60 m de comprimento. Na seção de testes, o escoamento é analisado por meio de sete estações de medição, cada uma equipada com um par de sensores de impedância elétrica, para acessar a concentração de fase, um sensor de temperatura e um sensor de pressão por estação.

Figura 14 – Diagrama do circuito experimental

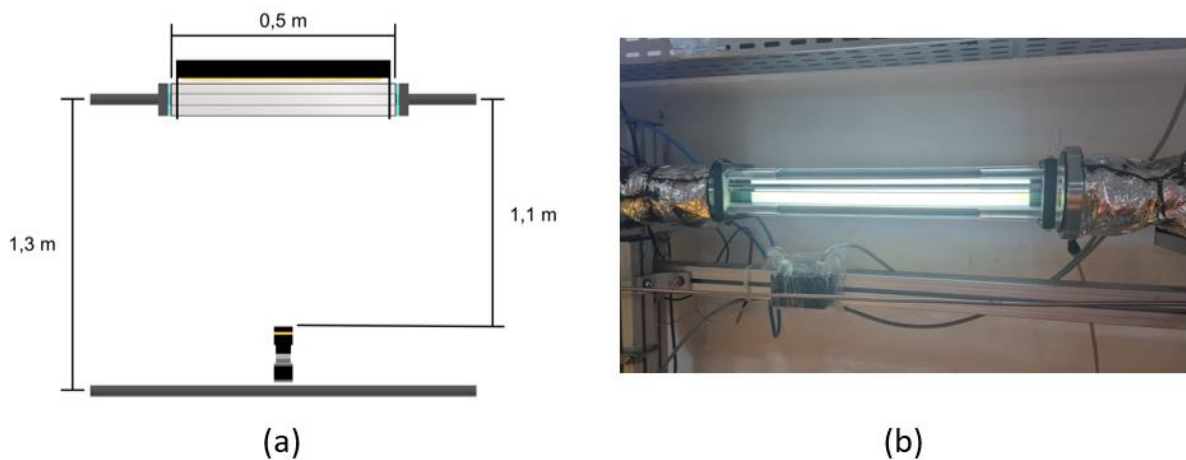


Fonte: Autoria própria.

Além das estações de medição, a seção de teste conta com uma estação de visualização no final do trecho experimental. A janela de visualização consiste em uma caixa acrílica transparente com seção quadrada de 7×7 cm e 50 cm de comprimento, contendo um orifício central de 25,4 mm de diâmetro, através do qual o escoamento passa e pode ser observado. A escolha de uma caixa acrílica com seção quadrada visa, primeiramente, minimizar as distorções óticas durante a aquisição de imagens, uma vez que superfícies planas reduzem efeitos de refração associados à curvatura de paredes cilíndricas. Além disso, esse formato confere robustez estrutural à janela de visualização e permite sua fixação por meio de flanges e barras roscadas, o que viabiliza a operação sob pressurização. O sistema de aquisição de imagens foi posicionado de forma centralizada em relação à janela de visualização. A Figura 15 apresenta: (a) um diagrama ilustrando o arranjo experimental, incluindo a câmera, a janela de visualização, o sistema de iluminação e suas respectivas distâncias; e (b) uma imagem da janela de visualização. A iluminação do sistema foi realizada por meio de fitas LED, com densidade de 240 LEDs por metro, potência de 20 W/m, temperatura de cor de 6500 K, tensão de operação de 12 V, ângulo de iluminação de 120° , índice de reprodução de cor (IRC) ≥ 80 e índice de proteção IP20.

A disposição da iluminação e câmera em lado opostos deve-se à utilização da técnica de retroiluminação (*shadowgraphy*). Essa técnica consiste em posicionar uma fonte de luz difusa atrás do tubo transparente, de modo que o feixe de luz atravesse o escoamento e seja capturado pela câmera do lado oposto. Como as fases escoando possuem diferentes índices de refração, as bolhas aparecem mais escuras, enquanto o líquido apresenta tonalidades mais claras na escala de cinza.

Figura 15 – Janela de visualização: (a) Esquemático; (b) Janela de visualização com iluminação ligada

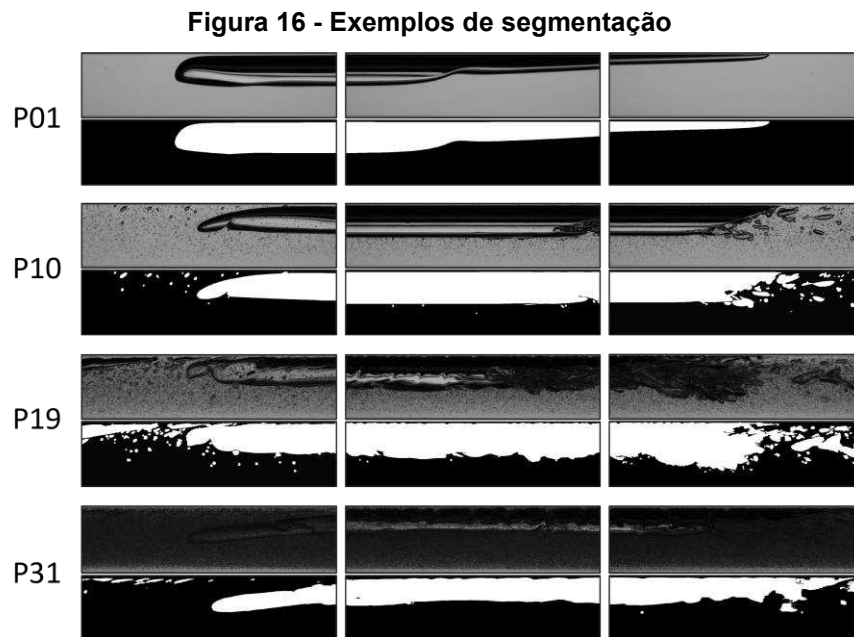


Fonte: Autoria própria.

Uma câmera Basler ace acA1440-220um, equipada com uma lente de 50 mm, foi utilizada para aquisição. A câmera foi configurada com uma taxa de aquisição foi de 500 Hz e tempo de exposição de 250 us. A câmera utilizada possui interface USB 3.0, e as imagens foram gravadas diretamente no computador, possibilitando o posterior processamento para extração dos parâmetros de interesse. As imagens em escala de cinza foram adquiridas com uma resolução de 1440 pixels de largura por 360 de altura, correspondendo a um *pixel size* de aproximadamente 0,08 mm, obtido pela razão entre o diâmetro interno do tubo e o número de pixels correspondente. Inicialmente as imagens incluíam margens horizontais adicionais para garantir que todo o diâmetro do tubo fosse enquadrado, mesmo em condições de pequenas vibrações. Posteriormente, essas margens foram removidas, resultando em uma imagem com proporção de 4:1 e dimensões aproximadas de 25,7 mm de altura por 102,8 mm de comprimento.

3.2 Segmentação de imagens

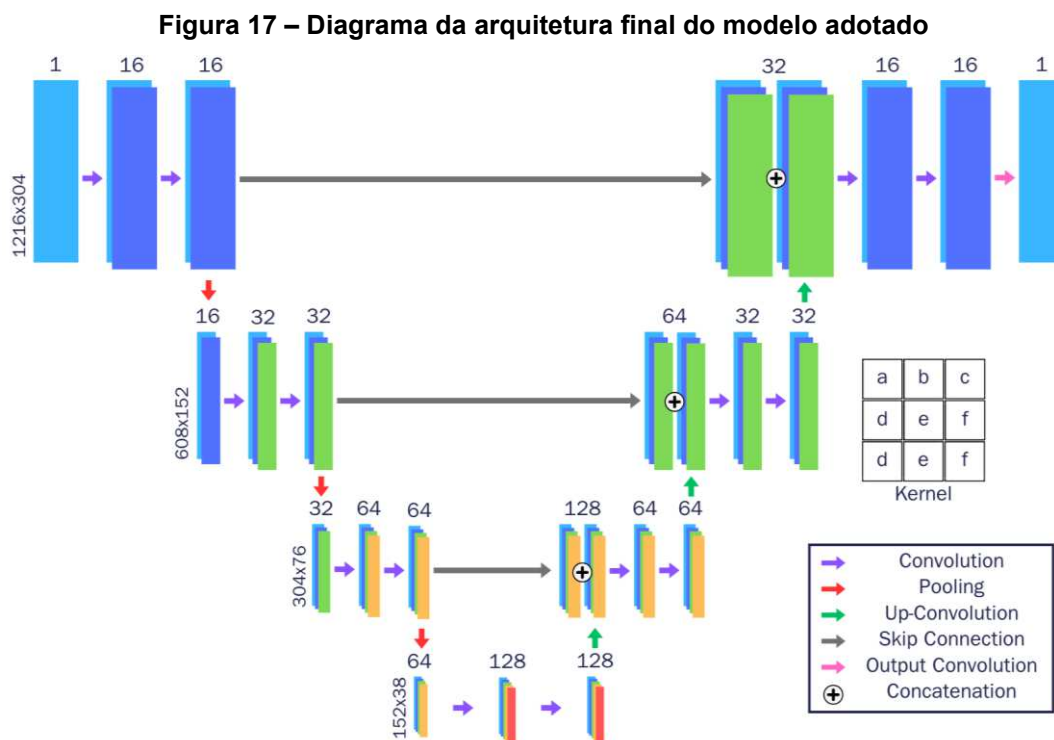
A segmentação de imagens constitui uma etapa central da metodologia proposta, pois fornece as máscaras binárias necessárias que servem de base para a identificação e rastreamento das bolhas. A Figura 16 apresenta exemplos de imagens em escala de cinza e suas respectivas segmentações de pontos experimentais da bateria de 4 bar. É possível observar nas imagens que a tarefa de segmentação se torna mais complexa devido a presença de distorções ópticas, ruído e artefatos visuais, como regiões claras no interior das bolhas, que podem ser erroneamente interpretadas como parte da fase líquida. Além disso, a presença de microbolhas pode alterar significativamente a intensidade média da imagem, tornando-a mais ou menos escura. Essa variação pode comprometer métodos dependentes de contraste bem definidos entre as fases. Nessas situações, as técnicas de aprendizado de máquina podem se destacar, devido a maior capacidade de generalização e adaptação a diferentes condições de imagem.



Fonte: Autoria própria.

A abordagem escolhida utiliza a arquitetura U-Net (Ronneberger; Fischer; Brox, 2015), amplamente reconhecida por sua eficiência e capacidade de generalização em tarefas de segmentação de imagens. Além de ser mais robusta a variações de contraste e iluminação, esse tipo de arquitetura requer uma quantidade relativamente reduzida de dados anotados para o treinamento, evitando o gargalo imposto pela necessidade de anotar manualmente muitas imagens. Um diagrama da

arquitetura utilizada neste trabalho é apresentado na Figura 17, indicando o número de camadas convolucionais, a quantidade de filtros e a resolução em cada nível da rede. O modelo recebe como entrada uma imagem de 1216x304 pixels que passa por uma primeira etapa de codificação, na qual a resolução espacial é progressivamente reduzida ao longo de 3 níveis, enquanto o número de filtros convolucionais é aumentado (16, 32 e 64). Após a última redução de resolução, são aplicadas duas camadas convolucionais de 128 filtros, correspondendo ao nível mais profundo e abstrato do modelo. Na etapa de decodificação, o modelo percorre o caminho inverso ao da codificação. A resolução é progressivamente aumentada ao longo de 3 níveis, enquanto o número de filtros em cada nível é reduzido (64, 32, 16). As informações correspondentes dos níveis de codificação são concatenadas às camadas da etapa de decodificação por meio das *skip connections*, de forma a evitar a perda de informação durante a reconstrução. Ao final da etapa de decodificação, a saída do modelo é a imagem segmentada, com dimensões idênticas as da imagem de entrada.

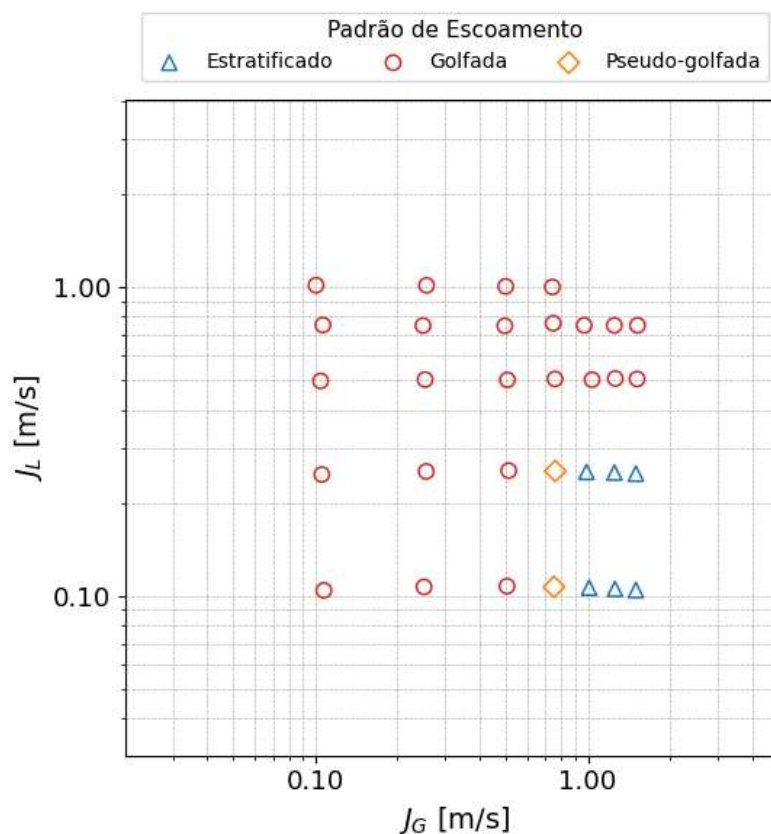


Fonte: Autoria própria.

Após definir a utilização da abordagem baseada em redes neurais convolucionais, foi necessário gerar um conjunto de dados representativo das diferentes condições experimentais observadas. Para isso, utilizou-se uma bateria de testes conduzidas a uma pressão na entrada da linha de 4 bar (bateria B1), composta

por 32 pontos experimentais a distintas velocidades superficiais de líquido e gás. Esses pontos, ilustrados no gráfico da Figura 18, correspondem a diferentes combinações de velocidades superficiais do líquido (J_L) e do gás (J_G). Embora o foco principal deste trabalho seja o escoamento em regime de golfadas, alguns pontos experimentais foram definidos em regiões próximas às fronteiras entre padrões de escoamento, levando à ocorrência de escoamentos estratificados e em regime de pseudo-golfadas em determinadas condições. Cada ponto experimental corresponde a 4 minutos de aquisição contínua de imagens, totalizando 120000 frames capturados a 500 Hz. Esse intervalo foi definido de forma a garantir a observação de um número estaticamente representativo de bolhas alongadas em caso de padrão intermitente, possibilitando uma análise dos parâmetros característicos do escoamento.

Figura 18 – Distribuição de velocidades superficiais



Fonte: Autoria própria.

Devido à alta frequência de aquisição, muitas imagens consecutivas apresentam aspectos visuais semelhantes, o que pode resultar em alta redundância temporal dos dados. Além disso, em alguns pontos experimentais, as bolhas alongadas excedem o campo de visão da câmera, fazendo com que a parte dos

quadros correspondam ao corpo da bolha, enquanto uma pequena fração representa o nariz, a cauda ou o pistão de líquido. Sabendo dessas duas características das imagens adquiridas, adotou-se uma estratégia baseada na separação das imagens em grupos de intensidade, seguida da realização de uma seleção aleatória proporcional dentro de cada grupo. Essa abordagem permitiu reduzir a redundância e abranger diferentes padrões visuais, resultando em um conjunto de dados mais representativo.

Inicialmente, as imagens de cada ponto experimental foram divididas em quatro faixas verticais de igual largura, ao longo da direção do escoamento. Em cada faixa, foi calculada a fração de pixels escuros, definida como a razão entre o número de pixels com intensidade inferior a 50 e o número total de pixels daquela região. Esse valor de limiar foi definido com base na análise visual dos histogramas de intensidade, de modo a englobar a maior parte dos pixels associados às bolhas, sem incluir pixels que pudessem gerar ambiguidades. Para cada ponto experimental e para cada faixa, determinaram-se os valores mínimo e máximo dessa fração ao longo de todas as imagens, os quais foram então quantizados em quatro níveis de ocupação de gás (0 até 3), igualmente espaçados entre os extremos observados. Dessa forma, cada imagem pôde ser associada a um vetor discreto de quatro níveis, representando o grau de presença de regiões escuras em cada faixa vertical, o que permite distinguir, por exemplo, estados sem bolha, bolha em entrada ou saída do campo de visão e bolha alongada ocupando praticamente todo o quadro.

A partir dessa codificação, as imagens foram agrupadas em classes definidas pela combinação dos níveis nas quatro faixas (grupos do tipo $1/1/2/3/4$) e, em seguida, foi realizada uma amostragem aleatória por faixa. Em cada grupo, foi selecionado um número limitado de imagens, utilizando todas as amostras disponíveis nos grupos raros e restringindo a quantidade máxima nos grupos muito abundantes, até atingir aproximadamente 3000 imagens por ponto experimental. Esse procedimento garante que diferentes estágios da passagem da bolha (sem gás, entrada, desenvolvimento e saída da bolha alongada) estejam representados de forma mais equilibrada no conjunto de dados utilizado para o treinamento do modelo de segmentação.

Uma vez concluído o processo de seleção das imagens representativas por ponto experimental, deu-se início à etapa de segmentação utilizando o modelo SAM. Inicialmente, as imagens foram recortadas de modo a eliminar as bordas horizontais adicionadas durante a aquisição, preservando apenas a região útil correspondente ao

diâmetro interno do tubo. Em seguida, com objetivo de suavizar ruídos e destacar o contraste entre as fases, foi aplicado um pré-processamento composto por inversão das intensidades, filtragem mediana e equalização adaptativa de histograma.

As imagens pré-processadas em escala de cinza foram convertidas para RGB e utilizadas como entrada para o modelo SAM, gerando máscaras binárias dos possíveis objetos (bolhas, fundo de líquido e estruturas residuais). Como modelo apresenta limitações na segmentação de regiões muito pequenas, as quais podem ser confundidas com ruído ou artefatos da imagem, regiões isoladas com área inferior a 50 pixels brancos foram descartadas. Para cada imagem de entrada, todas as máscaras binárias geradas foram salvas, juntamente com imagem pré-processada utilizada na inferência. Esse procedimento forneceu, de forma padronizada, o contorno individual dos objetos presentes nas imagens, os quais foram utilizados para compor uma máscara final, representando a fase gasosa em branco. Para identificar se as máscaras geradas pelo SAM correspondiam a regiões ocupadas por bolhas, foi implementado um critério baseado na intensidade média da imagem em escala de cinza, calculada na área onde a máscara apresentava pixels brancos. As máscaras cujas regiões apresentaram intensidade média inferior a 50 foram consideradas representativas de bolhas, uma vez que as bolhas aparecem mais escuras sobre o fundo claro das imagens. Esse critério evita a segmentação de efeitos de iluminação ou do fundo líquido como regiões brancas, mesmo quando identificados pelo SAM como objetos.

Em seguida, as máscaras que atenderam todos os critérios de corte previamente definidos foram somadas em uma máscara composta, sobre o qual se verificou o percentual de cobertura das regiões mais escuras da imagem em escala de cinza. Um valor elevado desse percentual indica que as principais regiões escuras haviam sido identificadas, classificando como uma boa candidata para o treinamento do modelo. Por fim, eventuais vazios internos nas regiões de bolhas foram preenchidos, evitando que a região com maior intensidade no centro das bolhas fosse interpretada como fase líquida. Após essa etapa, realizou-se uma análise visual final para remover os casos em que a máscara final não correspondia adequadamente ao contorno das bolhas escuras nas imagens em escala de cinza. Todo o procedimento resultou em um conjunto de dados final composto por aproximadamente 28500 pares de imagens, cada um contendo a imagem original em escala de cinza e sua respectiva máscara binária, utilizadas no treinamento do modelo.

Após a geração do conjunto de dados semiautomático utilizando o modelo SAM, tornou-se necessário conduzir uma série de experimentos para definir tanto a profundidade adequada da arquitetura U-Net quanto os critérios de seleção das imagens empregadas no treinamento. Uma análise exploratória preliminar foi realizada para avaliar o impacto do tamanho do conjunto de dados, variando desde 250 até aproximadamente 20000 pares de imagens, com estratificação por ponto experimental e por intensidade luminosa. A avaliação foi conduzida por meio de validação interna (treino/validação) e por dois conjuntos externos de teste: um conjunto amplo de aproximadamente 4500 imagens, com máscaras geradas automaticamente pelo SAM, e um conjunto reduzido de 80 imagens, composto por máscaras anotadas manualmente, utilizado como referência de alta confiabilidade. Dessa forma, buscou-se quantificar como o número de amostras influencia o desempenho da segmentação e identificar o ponto a partir do qual o ganho marginal decorrente do aumento do tamanho do conjunto de imagens se torna pouco significativo.

Além disso, foi estabelecido um protocolo experimental consistente para garantir a reprodutibilidade e a comparabilidade dos resultados. Foram conduzidos cinco subconjuntos de validação independentes, cada um inicializado com uma semente aleatória distinta, e para cada subconjunto foram treinados modelos utilizando 10 diferentes quantidades de pares de imagens para treinamento. Todos os treinamentos empregaram a mesma arquitetura U-Net e utilizaram mecanismos automáticos de ajuste e estabilização do aprendizado, garantindo consistência entre as diferentes execuções. A divisão do conjunto de treinamento foi realizada de forma estratificada por ponto experimental e intensidade luminosa, assegurando que todas as condições relevantes estivessem representadas de maneira equilibrada. As avaliações finais foram sempre realizadas sobre os mesmos conjuntos externos de teste semiautomático e manual, garantindo coerência na comparação entre os modelos obtidos.

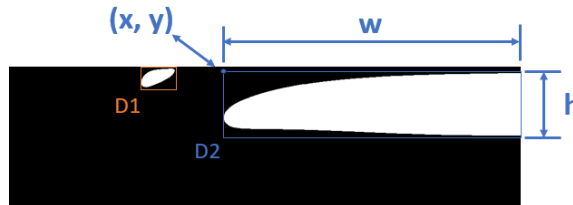
3.3 Rastreamento de objetos

A partir das imagens segmentadas pelo modelo de aprendizado profundo, é possível identificar, em cada frame, as regiões correspondentes a cada bolha presente no escoamento. O próximo passo consiste em associar as aparições individuais ao longo da sequência temporal, permitindo que cada bolha seja identificada e seguida

de maneira contínua. Para isso, foi desenvolvido um sistema de rastreamento de objetos baseado em detecção, no qual as correspondências são determinadas a partir de modelos de predição e associação. O sistema de rastreamento proposto foi desenvolvido com base nos modelos SORT e DeepSORT, que combinam o Filtro de Kalman e o Algoritmo Húngaro para realizar a associação ótima entre predições e detecções. A partir dessa estrutura, foram implementadas modificações na função de custo e adicionados mecanismos de rejeição (*gates*) destinados a eliminar associações incorretas, aprimorando a robustez do rastreamento em condições de ruído e oclusões parciais.

No sistema desenvolvido, cada detecção corresponde a identificação individual de uma bolha na máscara binária. Cada detecção é representada por uma caixa delimitadora (x, y, w, h) , que define a posição do canto superior esquerdo e as dimensões. A partir dessas detecções, o algoritmo mantém um conjunto de trilhas ativas (T), cada uma associada a uma bolha identificada. As trilhas armazenam o histórico temporal de cada bolha, incluindo a posição e velocidade do centroide, as dimensões e a área, além de um identificador numérico associado à trilha e o índice dos frames de onde cada informação foi extraída. A Figura 19a apresenta um frame contendo as detecções identificadas e, na Figura 19b, uma tabela representando um conjunto de trilhas. Os centroides (cx, cy) são calculados a partir do centro geométrico das caixas delimitadoras, já a velocidade é determinada com base no deslocamento do centroide entre frames.

Figura 19 – (a) Frame segmentado com detecções e caixas delimitadoras; (b) Matriz representando o conjunto de trilhas associadas entre frames sucessivos



(a)

ID	Frame	cx	cy	vx	vy	w	h
2	1	1156	152	—	—	80	45
2	2	1115	151	-41	-1	120	50
2	3	1074	153	-41	+2	180	54
2	4	1035	154	-39	+1	240	56
2	5	995	152	-40	-2	300	58
2	6	955	151	-40	-1	360	60
2	7	915	152	-40	+1	400	60

(b)

Fonte: Autoria própria.

A estrutura geral do rastreador proposto é composta por etapas sequenciais de detecção, predição, associação, atualização e gerenciamento de trilhas. Inicialmente, no frame analisado, todos os objetos são identificados por meio das caixas delimitadoras. Em seguida, o movimento é predito para todas as trilhas ativas no frame anterior, utilizando o modelo de movimento do Filtro de Kalman. Com as predições obtidas, verifica-se se há detecções que correspondem às trilhas existentes, por meio da etapa de associação implementada com o algoritmo Húngaro. As associações são resolvidas de forma global, minimizando o somatório total da função de custo. Em seguida, cada trilha é atualizada com base nas medições associadas e nas predições correspondentes, através do ganho de Kalman. Por fim, a etapa de gerenciamento de trilhas define quais trilhas permanecem ativas para análise nos frames seguintes e quais devem ser criadas ou encerradas.

O rastreador proposto adota um vetor de estado definido como $x = [cx, cy, vx, vy, w, h]^T$, sendo composto pelas coordenadas do centroide componentes de velocidade e dimensões da caixa delimitadora. A predição da posição no frame seguinte é obtida a partir do modelo de movimento definido pelas Equações 17 e 18, considerando que as variações de velocidade são tratadas como ruído de processo.

Para associação entre as predições e novas detecções, foi implementada a matriz de custo C_{ij} , na qual cada elemento representa o custo de associar a predição i com a detecção j . A função de custo é apresentada na Equação 20, a função combina métricas de distância e similaridade espacial. A distância de Mahalanobis normalizada é utilizada como métrica de dissimilaridade, de modo que valores mais altos indicam maior custo de associação entre a predição e a detecção. Já o termo $(1-IoU)$ atua de forma complementar, favorecendo pares com maior sobreposição entre as caixas delimitadoras e, portanto, maior similaridade espacial. O parâmetro λ , define o equilíbrio entre os componentes de distância e similaridade espacial.

$$C_{ij} = \lambda \frac{d_M(i, j)}{d_{M_{max}}} + (1 - \lambda)(1 - IoU(i, j)) \quad (20)$$

Onde:

- C_{ij} é o custo da associação i e a detecção j ;
- λ é o parâmetro de ponderação responsável por ajustar a contribuição relativa entre o termo de distância e o termo de similaridade espacial ($0 \leq \lambda \leq 1$);
- $d_M(i, j)$ é a distância de Mahalanobis entre a predição i e a detecção j ;
- $d_{M_{max}}$ é o valor máximo utilizado na normalização;
- $IoU(i, j)$ é o índice de intersecção sobre união.

No rastreador proposto, os pares predição-detecção são filtrados antes da otimização pelo algoritmo Húngaro, reduzindo o número de falsas correspondências e otimizando o uso dos recursos computacionais. Emprega-se os seguintes limiares de rejeição:

- Limiar de Mahalanobis: rejeita pares cujo desvio estatístico excede o valor máximo permitido ($d_M > 4,2$);
- Limiar vertical: descarta associações com deslocamento vertical superior a 25 px, considerado fisicamente improvável para o movimento das bolhas;
- Limiar mínimo de IoU: elimina pares com sobreposição entre caixas delimitadoras inferior a 0,7;

- Limiar direcional: rejeita pares com movimento incompatível com a direção predominante do escoamento (da direita para a esquerda).

Na etapa final de gerenciamento de trilhas, detecções não associadas servem de base para novas trilhas, que só são confirmadas após um número mínimo de atualizações consecutivas. As trilhas que permanecem sem detecção por até 10 frames consecutivos são mantidas ativas por predição. Ultrapassado esse limite, são encerradas, removendo-se as atualizações mais recentes que não apresentaram correspondência com detecções observadas. Essa abordagem assegura que bolhas identificadas de forma consistente ao longo de vários frames resultem em trilhas longas e estáveis, enquanto detecções ruidosas, fragmentadas ou decorrentes de artefatos e variações de iluminação não atendem aos critérios de confirmação e são automaticamente descartadas. Para cada identificação, são armazenadas as informações referentes à predição, à detecção e ao vetor de estado final após a etapa de correção, além do identificador de rastreamento (*ID*) e dos índices dos frames analisados.

3.4 Parâmetros de escoamento

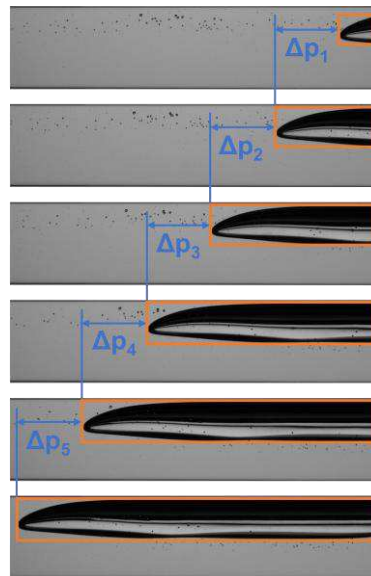
A determinação dos parâmetros característicos do escoamento foi conduzida baseando-se na célula unitária como referência estrutural. O cálculo dos parâmetros foi organizado em duas etapas complementares: (i) parâmetros derivados diretamente do rastreamento temporal das bolhas e (ii) parâmetros obtidos a partir das imagens agregadas correspondentes a cada célula unitária identificada e reconstruída.

Inicialmente, as bolhas identificadas na etapa de rastreamento foram filtradas de modo a selecionar apenas aquelas classificadas como bolhas alongadas. Essa classificação foi realizada quadro a quadro das detecções de cada bolha ao longo do tempo, com base em três critérios geométricos principais, avaliados ao longo das diferentes aparições de cada bolha nos frames: comprimento superior a $2D$, raio máximo maior que $D/3$ e área superior a 65000 pixels. Esses parâmetros foram definidos a partir da análise dos pontos experimentais que apresentavam bolhas menores e mais críticas para a classificação, nos quais se avaliou quais detecções deveriam ser aceitas ou rejeitadas pelo algoritmo através das imagens. Embora dependentes das características do conjunto de dados utilizado, tais valores são

facilmente ajustáveis e mostraram-se adequados para classificação das bolhas alongadas para os pontos analisados. Considerou-se como bolha alongada toda aquela que, em pelo menos uma de suas aparições, apresentou simultaneamente essas características geométricas. Os limiares adotados foram definidos a partir da análise estatística de bolhas não alongadas, de forma a distinguir de maneira consistente as estruturas de gás que apresentam morfologia típica de bolhas alongadas. Após a filtragem das bolhas alongadas, foi criado um arquivo de referência contendo, para cada célula unitária, o ID correspondente à fase de rastreamento e um novo identificador exclusivo (ID da célula unitária), além do primeiro frame de ocorrência. Esse arquivo foi estruturado para armazenar os parâmetros calculados por célula unitária, com cada coluna representando um parâmetro distinto.

Na primeira etapa, foram calculadas as velocidades médias das bolhas e a frequência de passagem, com base nos resultados obtidos na etapa de rastreamento de objetos. A velocidade da bolha foi definida como a média da velocidade do nariz da bolha ao longo de sua trajetória. Para essa estimativa, foi analisado o deslocamento do lado esquerdo das caixas delimitadoras entre frames consecutivos, correspondendo à posição do nariz da bolha. Foram desconsiderados os casos em que o deslocamento era nulo, situação que ocorre quando a bolha entra em contato com a borda esquerda do campo de visão. A Figura 20 ilustra, por meio de caixas delimitadoras, a posição do nariz da bolha em frames consecutivos.

Figura 20 - Deslocamento do nariz da bolha em frames consecutivos



Fonte: Autoria própria.

A velocidade média da bolha alongada, pode ser então calculada por meio da Equação 21, a partir da média dos deslocamentos do nariz da bolha entre frames consecutivos, multiplicada pelo tamanho de pixel e pela frequência de aquisição.

$$U_B = \frac{S_p f_{ac}}{N} \sum_{i=1}^N \Delta p_i \quad (21)$$

Onde:

Ainda na primeira etapa, duas frequências distintas foram determinadas a partir das sequências de células unitárias identificadas: a frequência média global (f_{mean}) e a frequência instantânea (f_{inst}). A frequência média global representa o número total de células unitárias que atravessam o campo de visão durante o intervalo de tempo de aquisição, sendo calculada por:

$$f_m = \frac{(N_b - 1)f_a}{N_f^{(1 > N_b)}} \quad (22)$$

Onde:

- N_b : número de bolhas alongadas detectadas;
- N_f : número de frames entre a primeira e a última detecção;
- f_a : frequência de aquisição.

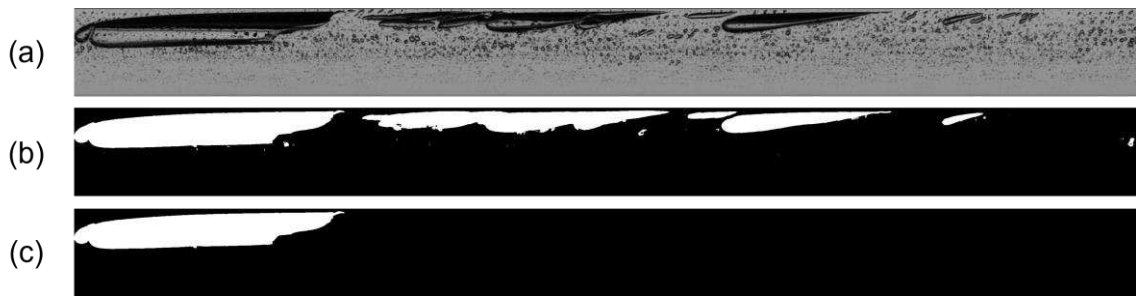
O termo $(N_{bubbles} - 1)$ é utilizado porque a última célula unitária observada não possui uma sucessora, impossibilitando a determinação do intervalo completo correspondente a um ciclo. Essa última célula é mantida apenas para verificação do término da penúltima célula, mas não é incluída no cálculo da frequência, uma vez que seu pistão de líquido não é totalmente visível dentro do campo de visão.

Já a frequência instantânea foi obtida para cada par de células unitárias consecutivas, a partir do intervalo de tempo entre o início de uma célula unitária e o início da célula seguinte, de acordo com:

$$f_{inst_i} = \frac{1}{N_f^{(i > i+1)}} \quad (23)$$

Com as bolhas alongadas e suas respectivas velocidades previamente identificadas, tornou-se possível gerar uma visualização representativa da célula unitária. A técnica empregada foi inspirada em Do Amaral et al. (2013) e consiste em agregar as porções finais das imagens sucessivas, deslocando-as conforme a velocidade média da bolha. Esse procedimento permite reconstruir a célula unitária completa, fornecendo uma representação espacial contínua do escoamento, mesmo quando a bolha não é inteiramente visível no campo de visão da câmera. Foram gerados três tipos de imagens agregadas, correspondendo a diferentes representações do escoamento: (a) imagem agregada das imagens em escala de cinza, (b) imagem agregada das máscaras binárias e (c) imagem agregada das máscaras binárias contendo apenas a bolha alongada. A Figura 21 apresenta esses três casos, exibidos em ordem para fins de comparação visual.

Figura 21 – Exemplo de imagens agregadas

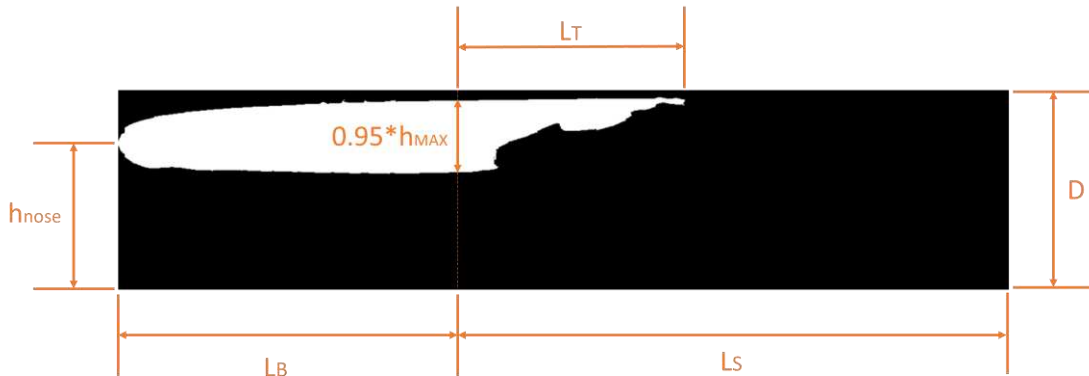


Fonte: Autoria própria.

Os parâmetros característicos relacionados a medidas geométricas da célula unitária são estimados a partir da imagem agregada das máscaras binárias contendo apenas a bolha alongada. Os comprimentos de bolha (L_b), do pistão de líquido (L_s) e a altura de nariz, são obtidos diretamente das imagens em número de pixels e, em seguida, convertidos para metros por meio da multiplicação pelo tamanho de pixel. O tamanho de pixel (s_p) foi definido como a razão entre o diâmetro interno da tubulação ($D = 25,4 \text{ mm}$) e o número de colunas da imagem (304 pixels), resultando em aproximadamente $0,08 \text{ mm/pixel}$. O final da bolha foi definido a partir da extremidade traseira da máscara (cauda) e determinado como a primeira posição, em direção ao centro da bolha, onde a altura local atinge 95% da altura máxima. Esse ponto marca a transição entre a região de cauda e o corpo principal da bolha, permitindo identificar de forma consistente o limite final da bolha alongada mesmo em casos com gradiente de espessura na cauda. A Figura 22 apresenta um exemplo de célula unitária

binarizada e a representação das medidas geométricas consideradas na sua caracterização.

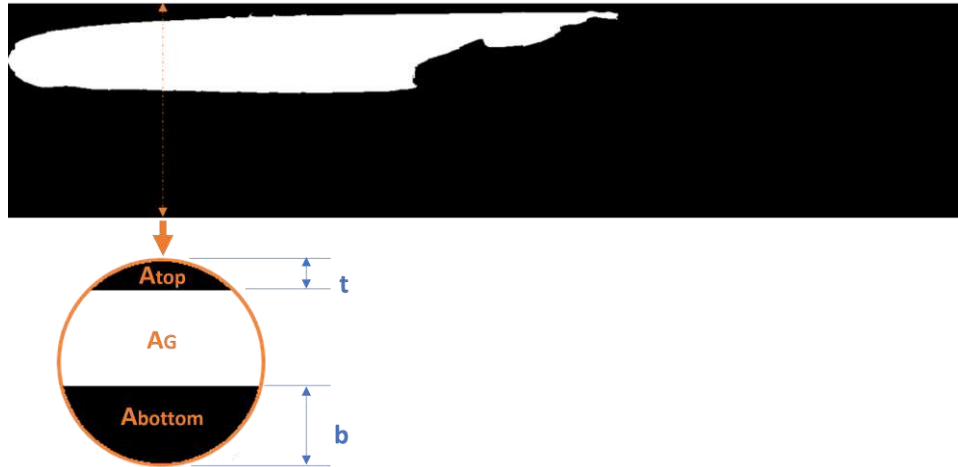
Figura 22 – Representação da célula unitária e medidas geométricas



Fonte: Autoria própria.

A fração de vazio (α) foi determinada a partir da imagem segmentada da célula unitária, considerando a distribuição da fase gasosa ao longo da tubulação. A análise foi realizada coluna a coluna na imagem agregada, correspondendo a seções transversais sucessivas do escoamento. Para cada coluna, foi calculada a fração de vazio local $\alpha(x)$, definida como a razão entre a área ocupada pelo gás e a área total da seção transversal da tubulação. A Figura 23 apresenta uma representação da seção transversal, onde A_G é a área de gás e a soma de A_{top} e A_{bottom} representa a área de líquido. A área total da seção transversal foi assumida como uma circunferência completa, enquanto a área de gás foi obtida a partir da região delimitada entre os dois semicírculos de líquido (regiões pretas da máscara binária). A fração de vazio na célula unitária foi calculado pela média das frações de vazio das seções transversais, sendo equivalente à razão entre o volume total de gás e o volume total da célula unitária.

Figura 23 – Representação da célula unitária e medidas geométricas



Fonte: Autoria própria.

As equações utilizadas para o cálculo da fração de vazio (α) da célula unitária são apresentadas nas Equações 24 a 27. A Equação 24 define a área da seção transversal circular (A_{seg}) compreendida entre a borda da tubulação e uma distância radial s , sendo aplicada no cálculo tanto de A_{bottom} quanto de A_{top} , por meio dos respectivos valores de b e t .

$$A_{seg}(s) = R^2 \arccos\left(1 - \frac{s}{R}\right) - (R - s)\sqrt{2Rs - s^2} \quad (24)$$

$$A_G = \pi R^2 - A_{top} - A_{bottom} = \pi R^2 - A_{seg}(t) - A_{seg}(b) \quad (25)$$

$$V_G = \sum_{i=1}^{N_{cols}} A_{G_i} s_p \quad (26)$$

$$\alpha = \frac{V_G}{V_T} = \frac{\sum_{i=1}^{N_{cols}} A_{G_i}}{\pi R^2 N_{cols}} \quad (27)$$

Onde:

- R é o raio da tubulação;
- V_G é o volume de gás na célula unitária;
- V_T é o volume total da célula unitária;
- N_{cols} é o número de colunas da imagem agregada.

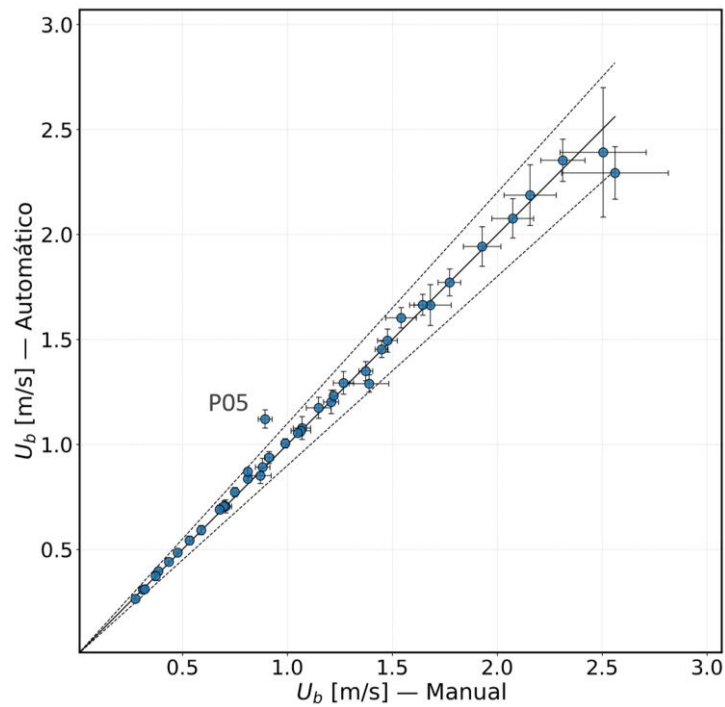
4 RESULTADOS

Este capítulo apresenta os principais resultados obtidos no desenvolvimento deste trabalho. Inicialmente, são avaliados o desempenho e a qualidade da metodologia de segmentação de imagens, etapa fundamental para a identificação e o rastreamento das bolhas. Nessa fase, são apresentadas as métricas de avaliação das etapas de treinamento e validação do modelo, utilizando tanto máscaras anotadas manualmente quanto máscaras geradas de forma semiautomática. Adicionalmente, são mostrados exemplos representativos de imagens segmentadas, para diferentes velocidades superficiais, com o objetivo de ilustrar o desempenho do método e evidenciar suas limitações. A título exploratório, também são realizados testes com imagens provenientes de experimentos fora do escopo principal deste trabalho, a fim de verificar a capacidade de generalização do modelo proposto.

Em seguida, são apresentados os resultados dos parâmetros característicos do escoamento bifásico, calculados a partir da análise das máscaras segmentadas. Essa análise contempla diferentes baterias experimentais, permitindo a avaliação a reprodutibilidade da metodologia para situações distintas. Adicionalmente, alguns dos parâmetros obtidos são comparados aos valores medidos por sensores capacitivos, de modo a verificar a consistência e confiabilidade dos resultados.

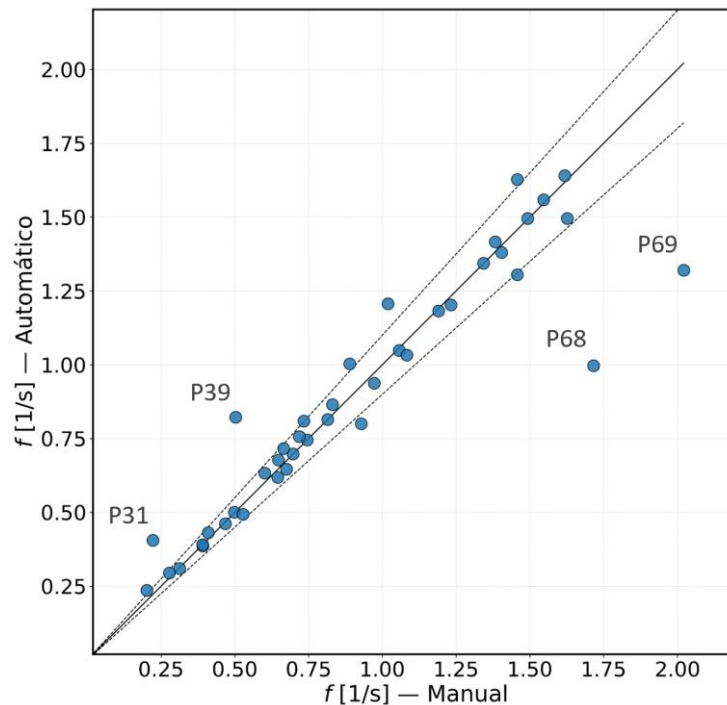
Inicialmente, como parte da análise exploratória preliminar, o modelo de segmentação foi treinado com 3500 imagens provenientes da bateria experimental de 3 bar, atingindo um coeficiente *Dice* médio da ordem de 0,925. Essa etapa teve como objetivo avaliar a viabilidade inicial do modelo de segmentação proposto. As comparações entre os resultados obtidos pelos métodos automático e manual de processamento de imagens, para os parâmetros de velocidade e frequência, são apresentadas nas Figuras 24 e 25.

Figura 24 – Velocidade de bolhas alongadas para bateria de 3 bar



Fonte: Autoria própria.

Figura 25 – Frequência de bolhas alongadas para bateria de 3 bar

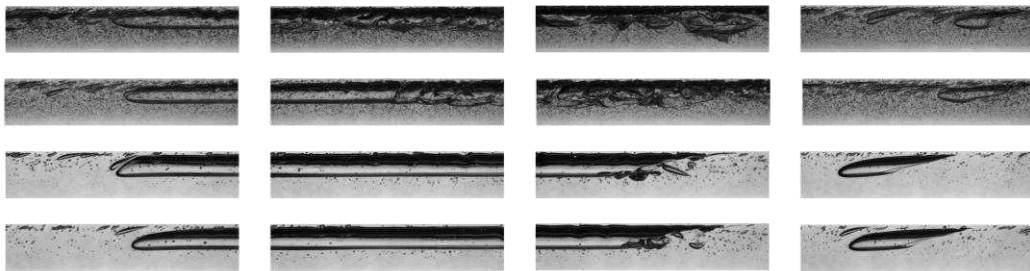


Fonte: Autoria própria.

A partir da análise dos gráficos para velocidade e frequência obtidos para condição de 3 bar, observa-se que os valores permanecem, em sua maioria em um intervalo de $\pm 10\%$ em relação às medidas manuais. Essa concordância é reforçada

pelos erros calculados, sendo obtidos, para a velocidade de bolha (U_B), RMSE de 0,06 m/s e MAPE de 2,7% e, para frequência (f), RMSE de 0,18 Hz e MAPE de 9,8%. O gráfico de velocidade apresenta um ponto fora deste intervalo, correspondente a uma condição em que apenas nove bolhas alongadas foram identificadas, como mostrado na Tabela 2, o que limita a representatividade estatística da velocidade média estimada. Já o gráfico de frequência, apresenta 4 pontos significativamente afastados do intervalo de 10%, os quais apresentam estruturas gasosas no limiar de serem ou não classificadas como bolhas alongadas, o que pode gerar ambiguidades na identificação e pode resultar em superestimação ou subestimação dos valores de frequência. A Figura 26 apresenta algumas imagens correspondentes a essas condições experimentais. Mesmo a presença desses pontos atípicos fora do intervalo de 10%, é possível observar que a grande maioria dos pontos apresenta boa concordância entre os resultados obtidos.

Figura 26 – Exemplos de imagens correspondentes aos pontos fora do intervalo de $\pm 10\%$ no gráfico de frequência (3 bar)



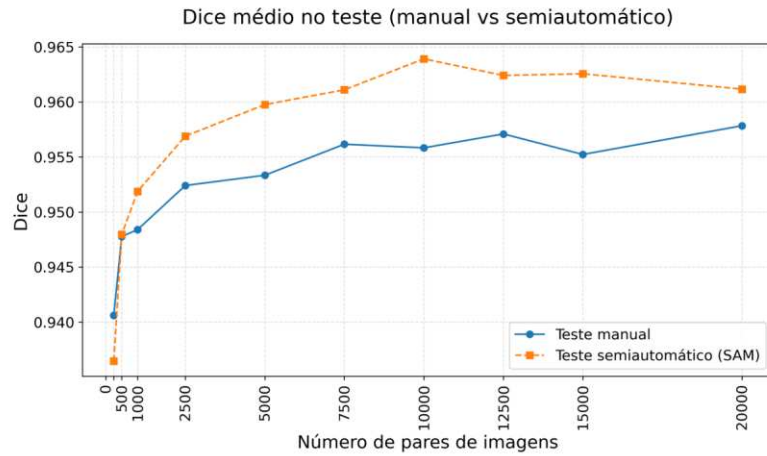
Fonte: Autoria própria.

4.1 Segmentação de imagens

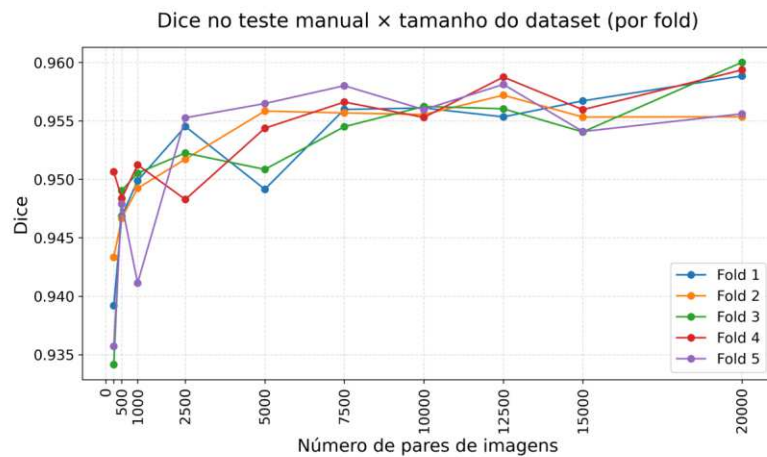
O modelo de segmentação final foi avaliado utilizando as três baterias experimentais a 4 bar. A avaliação da qualidade das máscaras segmentadas também foi conduzida através do coeficiente *Dice*, calculado separadamente para as máscaras anotadas manualmente e as máscaras geradas de maneira semiautomática pelo SAM. A Figura 27 (a) apresenta o coeficiente Dice médio para os testes utilizando as máscaras manuais e semiautomáticas. Esse gráfico demonstra o desempenho geral do modelo e evidencia a tendência da métrica de avaliação do modelo em relação com o aumento do conjunto de dados de treinamento. As Figuras 27 (b) e (c) apresentam os resultados individuais para cada partição da validação cruzada, considerando, respectivamente, as avaliações semiautomáticas e manuais. Esses gráficos permitem identificar eventuais diferenças entre partições, fazendo com que

seja possível analisar a estabilidade do modelo diante das distintas divisões do conjunto de treinamento.

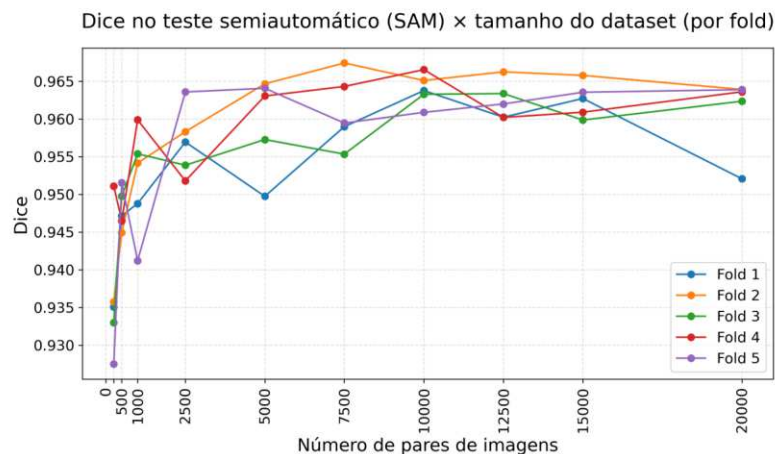
Figura 27 – Teste com diferentes conjuntos de dados. (a) Dice médio para máscaras manuais e semiautomáticas em função do tamanho do conjunto de dados. (b) Dice por subconjunto para validação cruzada das máscaras semiautomáticas. (c) Dice por subconjunto para validação cruzada das máscaras anotadas manualmente



(a)



(b)

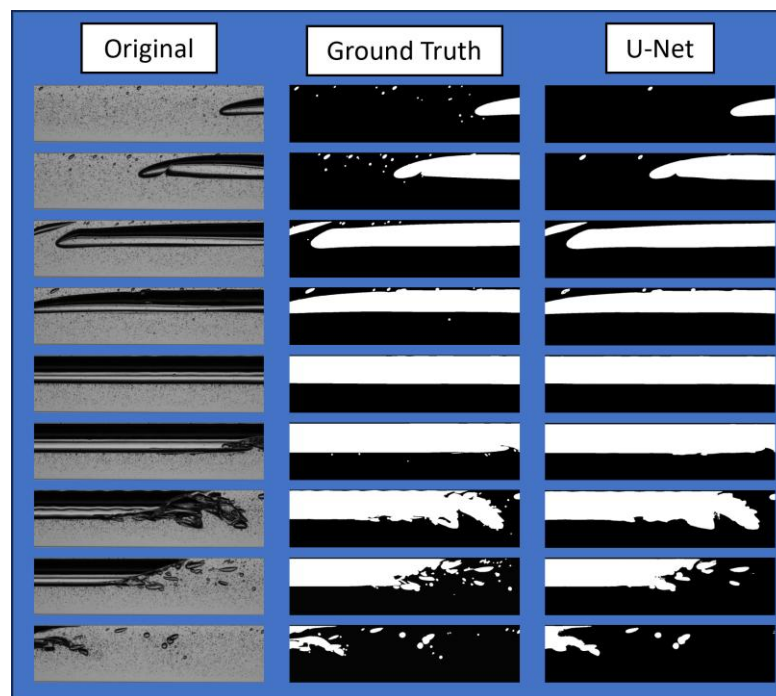


(c)

Fonte: Autoria própria.

A Figura 28 apresenta uma comparação entre a imagem original, a máscara de referência anotada manualmente (*ground truth*) e a segmentação gerada com o modelo U-Net. É possível observar algumas diferenças entre as imagens de referência e as imagens geradas com o modelo treinado, bolhas de menor dimensão tendem a serem identificadas como líquido. Esse comportamento já era esperado, devido às limitações impostas pelo contraste local e pela própria estratégia de treinamento. No caso específico das bolhas pequenas, a baixa definição e fraco contraste tornam sua identificação mais complexa, dificultando inclusive a anotação manual. Outra explicação para o menor desempenho nestes casos, está associada ao treinamento, dado que objetos com área inferior a 50 pixels foram excluídos do processo de treinamento.

Figura 28 – Comparação entre a imagem original, a máscara de referência (*ground truth*) e a máscara segmentada pela U-Net.



Fonte: Autoria própria.

Por outro lado, é possível perceber que as bolhas maiores são segmentadas com elevada consistência, apresentando contornos bem definidos mesmo para diferentes situações. Nos eventos de desprendimento ou coalescência de bolhas, nota-se que o modelo tende a agrupar bolhas muito próximas, resultando em uma estrutura única segmentada. Apesar dessas limitações nas regiões mais complexas, às métricas associadas a segmentação permanecem robustas, reforçando que o modelo é adequado para caracterização das bolhas alongadas.

4.2 Parâmetros de escoamento

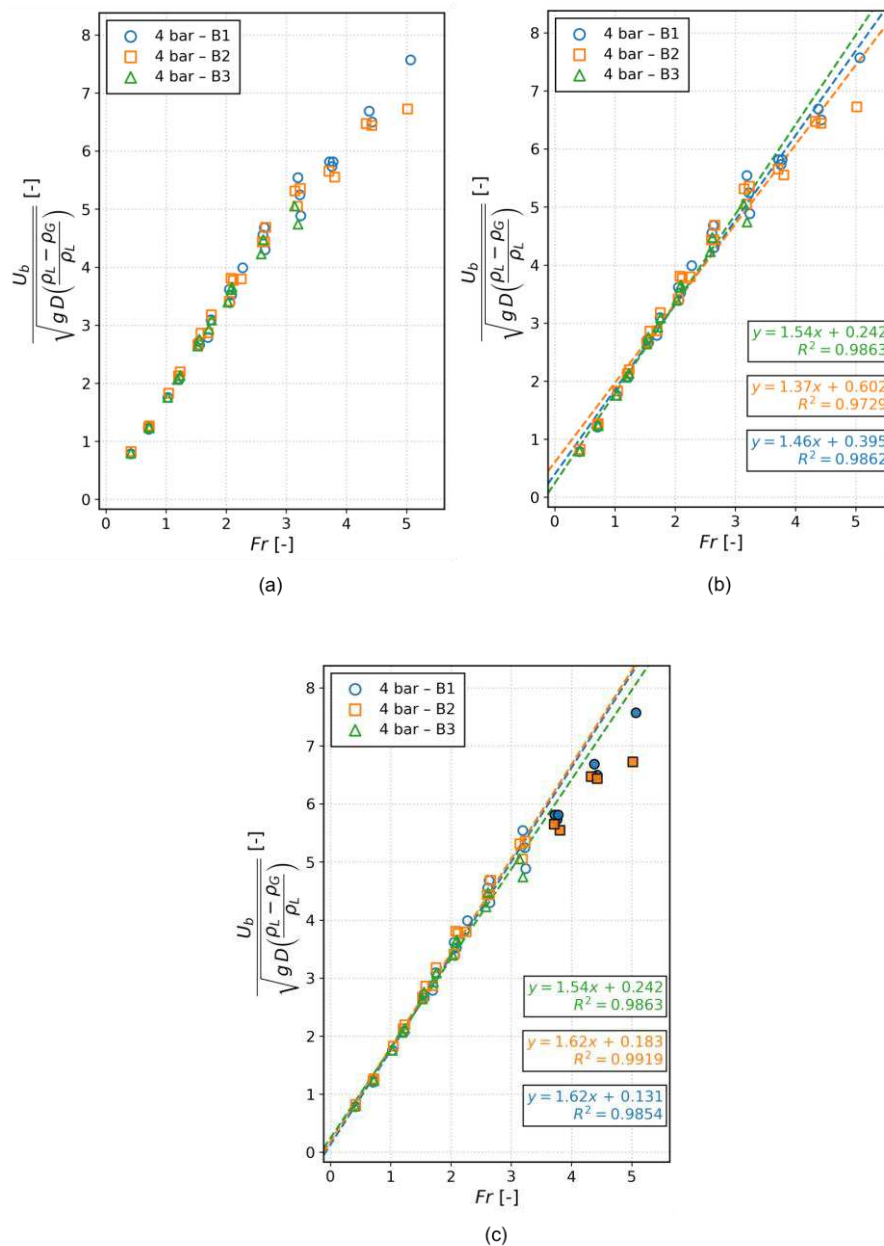
Nesta seção, são apresentados os resultados obtidos para os parâmetros que caracterizam o escoamento. Para cada ponto experimental, foi gerada uma tabela contendo todos os parâmetros calculados individualmente para cada célula unitária identificada. Foram analisadas três baterias experimentais, todas conduzidas à pressão de 4 bar. Além disso, foi construída uma tabela com parâmetros médios por bateria, possibilitando uma análise comparativa para as diferentes velocidades superficiais. Os principais parâmetros reportados incluem a velocidade de bolha alongada, frequência, fração de gás e comprimentos de célula unitária, bolha e pistão de líquido. Além disso, também são apresentados histogramas de todos esses parâmetros, permitindo visualizar a distribuição estatística.

Com o objetivo de avaliar as consistências dos parâmetros obtidos, os resultados foram comparados com modelos consolidados na literatura para escoamentos intermitente. Essa comparação atua como uma validação inicial da metodologia, possibilitando verificar se as grandezas extraídas estão compatíveis com o comportamento descrito por correlações clássicas. O primeiro parâmetro a ser verificado é a velocidade de bolha, que desempenha papel central em toda metodologia. A partir deste parâmetro, os procedimentos de agregação são conduzidos, influenciando na determinação dos demais parâmetros da célula unitária. Como apresentado na Equação 7, a comparação dos dados experimentais de velocidade de bolha alongada pode ser realizada através de uma análise adimensional. As velocidades calculadas através dos dados experimentais foram normalizadas conforme a expressão apresentada. A Figura 29 apresenta o resultado dessa análise, o eixo horizontal representa o número de *Froude* ajustado e o eixo vertical corresponde a velocidade normalizada calculada a partir dos dados experimentais. Adicionalmente, a análise comparativa entre as baterias experimentais evidenciou boa repetibilidade, expressa por uma diferença relativa média de 3,8% entre B1 e B2 e por um coeficiente de variação de 2,2% considerando os pontos experimentais em comum das baterias B1, B2 e B3.

Na subfigura (a) são apresentados os dados experimentais para todas as baterias experimentais investigadas. As subfiguras (b) e (c) apresentam os ajustes lineares aplicados aos dados experimentais, que podem ser adequadamente representados por uma equação do tipo $ax + b$. Os gráficos correspondem a três

baterias experimentais distintas e indicam boa reprodutibilidade do comportamento observado. No gráfico (c), onde os valores de Fr superiores a 3,5 são desconsiderados, observa-se uma melhora significativa na qualidade do ajuste, com coeficiente de R^2 igual a aproximadamente 0,99. Esse resultado indica que o sistema desenvolvido apresenta maior dificuldade para regimes caracterizados por números de Froude elevados, os quais estão associados a maiores velocidades superficiais e, portanto, condições mais complexas. Ainda assim, mesmo nesses casos, o método é capaz de fornecer resultados coerentes.

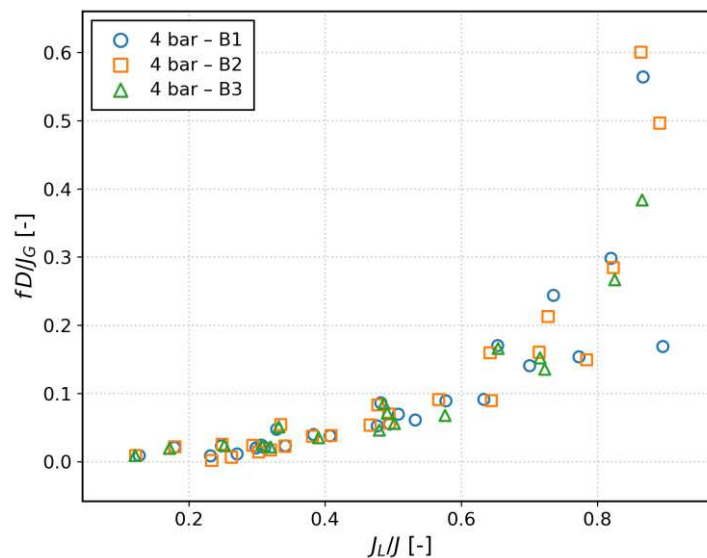
Figura 29 – Modelo clássico de velocidade de bolha alongada



Fonte: Autoria própria.

Além da velocidade de bolhas alongadas, outro parâmetro fundamental para análise de escoamentos intermitentes é a frequência de passagem de células unitárias. Neste trabalho, a análise da frequência é realizada seguindo uma metodologia semelhante a proposta por Fossa, Guglielmini e Marchitto (2003), onde os dados experimentais são avaliados através do número de Strouhal em função da fração de velocidade superficial do líquido. O gráfico para as três baterias experimentais é apresentado na Figura 30, observa-se um comportamento não linear, com uma tendência crescente, onde o número de Strouhal assume valores reduzidos para baixos valores de J_L/J e valores acentuados a partir de um determinado limiar. Esse comportamento é consistente com o reportado na literatura para regimes intermitentes, associando um aumento da taxa de formação de bolhas alongadas à medida que a fração de velocidade superficial do líquido aumenta. Neste caso, obteve-se uma diferença relativa média de 17,6% entre as baterias B1 e B2, e coeficiente de variação de 7,8% para B1, B2 e B3, considerando os pontos em comum.

Figura 30 – Número de Strouhal em função da fração de velocidade superficial do líquido

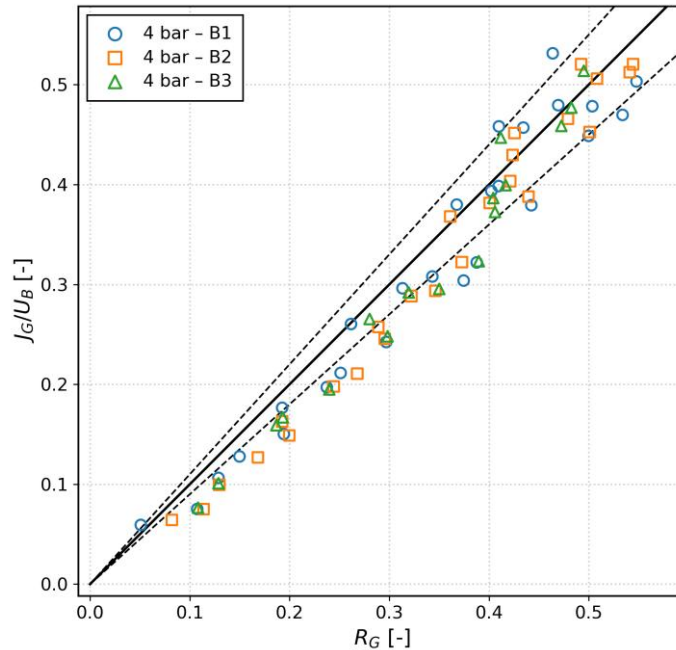


Fonte: Autoria própria.

Por fim, outra grandeza de interesse que pode ser comparada com estimativas da literatura é a fração de vazio, obtida experimentalmente através das imagens agregadas. Uma estimativa amplamente utilizada na literatura é baseada na razão entre a velocidade superficial de gás e a velocidade da bolha alongada. A comparação entre a fração de vazio experimental e a estimativa J_G/U_B permite avaliar a consistência dos valores obtidos e a coerência da metodologia de processamento de

imagem. A Figura 31 apresenta a comparação entre R_G (fração de vazio experimental) e a estimativa utilizando valores experimentais J_G/U_B , onde as linhas tracejadas pretas indicam os limites superior e inferior correspondentes a $\pm 10\%$ do valor estimado.

Figura 31 – Estimativa de fração de vazio em função da fração de vazio estimada experimentalmente



Fonte: Autoria própria.

Quantitativamente, obteve-se diferença relativa média de 3,1% entre B1 e B2, e coeficiente de variação médio de 1,9% para o conjunto B1, B2 e B3. Os valores de RMSE e MAPE obtidos para a estimativa da fração de vazio são apresentados na Tabela 3. É possível observar que os valores medidos apresentam boa coerência entre as baterias experimentais, indicando boa reprodutibilidade dos dados obtidos. Entretanto, na maioria das frações de vazio, os valores medidos tendem a ser ligeiramente superestimados. Esse comportamento decorre da metodologia adotada, que por se tratar de uma abordagem simplificada, considera toda a região entre a base e o topo da bolha em uma região transversal de uma bolha alongada como sendo ocupada por gás. No entanto, sabe-se que nessa região existe um filme líquido que envolve a bolha, o qual não é considerado no cálculo. Como as imagens utilizadas correspondem a uma visão lateral da bolha, não é possível obter diretamente informações completas da seção transversal do escoamento. Ainda assim, mesmo nesses casos, o método é capaz de fornecer resultados coerentes.

Tabela 3 – Métricas de erro da estimativa de fração de vazio

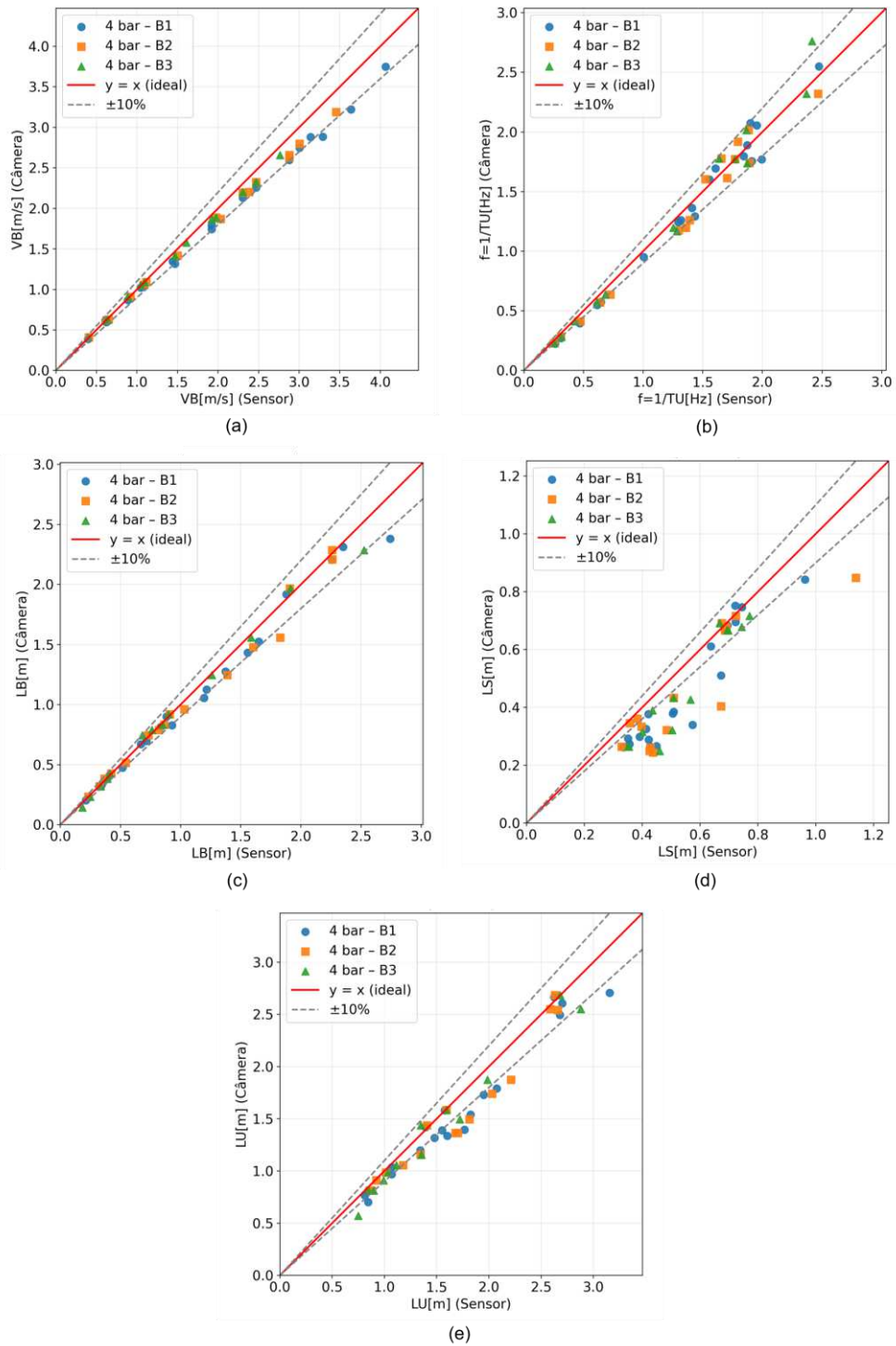
Bateria	Pontos	RMSE [-]	MAPE [%]
B1	26	0,04	13,50
B2	25	0,04	15,38
B3	17	0,03	13,72

Fonte: Autoria própria.

Adicionalmente, foi realizada uma verificação adicional por meio de uma metodologia independente baseada em sensores capacitivos de haste. Os dados foram adquiridos a 1000 Hz na estação de medição mais próxima da janela de visualização, utilizando sensores capacitivos dispostos em pares. A resposta dos sensores foi previamente caracterizada e, em seguida, convertida em séries temporais de nível de líquido. Com base nessas séries temporais, os parâmetros do escoamento foram estimados, incluindo a velocidade das bolhas, determinada a partir do atraso temporal entre os sinais dos sensores, e a frequência, calculada a partir da contagem de células unitárias identificadas no sinal após a aplicação de um limiar. Os parâmetros relacionados as dimensões da célula unitária também foram estimadas, com base na velocidade estimada anteriormente e nos sinais correspondentes a cada célula unitária.

Os parâmetros estimados por câmera e por sensor capacitivo são apresentados na Figura 32. Observa-se, nos gráficos de (a) velocidade e (b) frequência, boa concordância entre os métodos de medição, com quase todos os pontos concentrados dentro do intervalo de $\pm 10\%$. Em relação aos gráficos relacionados a comprimentos, observa-se boa concordância para o comprimento da bolha (c) e comprimento de célula unitária (e), enquanto o comprimento de pistão de líquido apresenta maior dispersão. Isso pode ser explicado pela dificuldade de definir com precisão o final da bolha alongada. Além do mais, como o pistão de líquido é geralmente muito menor que a bolha alongada, pequenas diferenças na sua identificação têm impacto significativo na sua estimativa.

Figura 32 – Comparação entre os parâmetros estimados por sensor capacitivo e por câmera



Fonte: Autoria própria

De forma complementar a Tabela 4 apresenta os valores de RMSE e MAPE obtidos para cada parâmetro da Figura 32, permitindo quantificar os desvios observados entre as medidas da câmera e as dos sensores capacitivos.

Tabela 4 - RMSE e MAPE obtidos para os parâmetros estimados pela câmera em comparação com os sensores capacitivos

Bateria	Número de Pontos	U_B		f		L_B		L_S		L_U	
		RMSE [m/s]	MAPE [%]	RMSE [Hz]	MAPE [%]	RMSE [m]	MAPE [%]	RMSE [m]	MAPE [%]	RMSE [m]	MAPE [%]
B1	19	0,21	7,3	0,1	7,9	0,11	5,5	0,11	27,2	0,2	10,8
B2	16	0,14	5,2	0,11	8,8	0,09	4,6	0,14	27,8	0,19	10
B3	13	0,07	2,7	0,12	7	0,07	6	0,1	25,1	0,14	9,4

Fonte: Autoria própria

O Apêndice B apresenta os histogramas dos principais parâmetros estimados neste trabalho, incluindo velocidade de bolha, fração de vazio e comprimentos de célula unitária, da bolha e do pistão, considerando as três baterias experimentais de 4 bar. Os gráficos estão organizados em uma malha relacionando as velocidades superficiais, sendo que cada célula contém a distribuição dos valores obtidos para cada ponto experimental correspondente. Esses histogramas permitem avaliar a dispersão e a consistência estatística dos parâmetros em diferentes condições de escoamento.

5 CONCLUSÃO

A análise dos resultados demonstra que a metodologia de processamento de imagens desenvolvida é capaz de estimar os parâmetros relevantes de escoamentos bifásicos gás-líquido em regimes em golfadas. A boa concordância observada com modelos consolidados, com medições manuais e com parâmetros estimados através de sensores capacitivos indica que a metodologia proposta é capaz de fornecer estimativas confiáveis dos parâmetros de interesse. As comparações dos parâmetros estimados apresentaram diferenças dentro de uma faixa de $\pm 10\%$ para a maior parte das condições experimentais. Os casos em que foram observadas maiores discrepâncias, geralmente estavam associados a condições de transição de regime ou bolhas no limiar de classificação como bolhas alongadas, o que introduz ambiguidades ao processo de identificação.

O modelo de segmentação U-Net, treinado com máscaras geradas pelo Segment Anything Model (SAM), apresentou valores de métricas de avaliação superiores a 0.95 para diferentes situações de escoamento. Essa abordagem permitiu reduzir significativamente o esforço de rotulagem manual, ao mesmo tempo que possibilitou a criação de um conjunto de dados grande e representativo. Além disso, o treinamento com diferentes tamanhos de conjuntos de dados, permitiu identificar o número de imagens necessário para alcançar um desempenho elevado de segmentação.

De forma geral, o método proposto se mostra uma alternativa promissora para a caracterização de escoamentos bifásicos gás-líquido. Entre as principais vantagens da metodologia destacam-se a visualização detalhada do escoamento e de sua reconstrução através do agregamento das imagens. Outra vantagem é a natureza não intrusiva da metodologia, embora sua aplicação em tubulações não transparentes dependa da instalação de uma janela de visualização. Uma limitação da metodologia está associada ao elevado volume de dados gerados, característica esperada de técnicas baseadas em aquisição de alta velocidade. Para a resolução e a frequência de aquisição utilizadas, um ponto experimental pode gerar 60 GB de imagens em escala de cinza. Essa quantidade de dados, associada a complexidade necessária para processamento de imagens, implica em tempos elevados de processamento, podendo atingir aproximadamente 3 horas por ponto experimental.

Os resultados apresentados abrem perspectiva para trabalhos futuros, relacionados tanto ao aprimoramento da metodologia proposta quanto à capacidade de generalização de modelos de segmentação baseados em redes convolucionais para escoamento bifásico. Em relação à estimativa dos parâmetros, percebe-se a necessidade de mudar a metodologia de cálculo da fração de vazio, reduzindo as simplificações adotadas e incorporando abordagens mais recentes, como apresentadas por Liu et al. (2022) e Perkins (2025), que buscam representar de forma mais realista a geometria da interface e a distribuição espacial das fases. Outro fator importante relacionado à estimação dos parâmetros, seria definir critérios mais robustos para identificação do término das bolhas alongadas e para definição de bolha alongada, ou pelo menos, tornar tais critérios facilmente controláveis. Embora uma única condição de pressão contemple diferentes condições experimentais de velocidades superficiais de líquido e de gás, seria relevante realizar testes adicionais em outras pressões, bem como com imagens provenientes de diferentes sistemas de aquisição de imagem. Nesse contexto, uma possível abordagem seria a padronização das imagens de entrada para uma resolução fixa, aliada à normalização dos histogramas de intensidade, tornando as imagens mais homogêneas mesmo para sistemas distintos de aquisição de imagens.

REFERÊNCIAS

ADAMS, Rolf; BISCHOF, Leanne. Correspondence seeded region growing. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 16, n. 6, 1994.

BARNEA, Dvora; SHOHAM, Ovadia; TAITEL, Yehuda. Flow pattern transition for downward inclined two phase flow; horizontal to vertical. **Chemical Engineering Science**, n. 5, p. 735–740, 1982.

BELTRÃO, R. L. C. *et al.* Challenges and new technologies for the development of the pre-salt cluster, Santos Basin, Brazil. *In*: Houston, Texas, USA: **Offshore Technology Conference**, maio 2009.

BENAMEUR, Narjes *et al.* An improved approach for cardiac MRI segmentation based on 3D UNet combined with papillary muscle exclusion. **Arabian Journal for Science and Engineering**, v. 50, n. 23, p. 19723–19741, 2024.

BENDIKSEN, Kjell H. An experimental investigation of the motion of long bubbles in inclined tubes. **International Journal of Multiphase Flow**, v. 10, n. 4, p. 467–483, 1984.

BERTOLA, Volfango. **Modelling and experimentation in two-phase flow**. Vienna: Springer-Verlag Wien, 2003.

BEWLEY, Alex *et al.* Simple online and realtime tracking. *In*: Phoenix: IEEE, 7 jul. 2017. Disponível em: <<http://arxiv.org/abs/1602.00763>>

BRILHADOR, Anderson; LAZZARETTI, André; LOPES, Heitor. A survey on open world learning for image segmentation: definitions, challenges, and directions. **Learning and Nonlinear Models**, v. 23, n. 1, p. 33–48, 27 jan. 2025.

CANNY, John. A computational approach to edge detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 8, n. 6, p. 679, 1986.

ÇENCEL, Yunus A.; CIMBALA, John M. **Mecânica dos fluidos - fundamentos e aplicações**. 2. ed. Porto Alegre: AMGHG Editora, 2012.

CHEN, Liang-Chieh *et al.* DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 40, n. 4, p. 834–848, 2 jun. 2016.

ÇIÇEK, Özgün *et al.* 3D U-Net: learning dense volumetric segmentation from sparse annotation. *In*: Cham: Springer, 21 jun. 2016. Disponível em: <<http://arxiv.org/abs/1606.06650>>

COELHO DE ANDRADE, Mauren Lousie Sguario. **Aplicação do método Level Set para segmentação e classificação de padrões e medidas de escoamento bifásico gás-líquido**. Curitiba: Universidade Tecnológica Federal do Paraná, 2011.

DA SILVA, Jaqueline Diniz *et al.* Development of a deep learning-based image processing technique for local phase fractions of multiphase transient flow. *In: Foz do Iguacu, PR, Brazil: ABCM, nov. 2024.*

DAS NEVES, Daniely A. *et al.* An experimental investigation on the fluid–structure coupling in horizontal pipes conveying two-phase intermittent flow. **International Journal of Multiphase Flow**, v. 176, p. 104825, 2024.

DAVIES, R.; TAYLOR, G. The mechanics of large bubbles rising through extended liquids and through liquids in tubes. **Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences**, v. 200, n. 1062, p. 375–390, 1949.

DE OLIVEIRA, W. R. *et al.* Bubble characterization in horizontal air-water intermittent flow. **International Journal of Multiphase Flow**, v. 69, p. 18–30, 1 mar. 2015.

DINH, T. Bui; KIM, B. S.; CHOI, T. S. Application of image processing techniques to air/water two-phase flow. *In: Denver, Colorado, USA: SPIE, 1999.*

DO AMARAL, C. E. F. *et al.* Image processing techniques for high-speed videometry in horizontal two-phase slug flows. **Flow Measurement and Instrumentation**, v. 33, p. 257–264, out. 2013a.

DO AMARAL, C. E. F. *et al.* Image processing techniques for high-speed videometry in horizontal two-phase slug flows. **Flow Measurement and Instrumentation**, v. 33, p. 257–264, out. 2013b.

FOSSA, M.; GUGLIELMINI, G.; MARCHITTO, A. Intermittent flow parameters from void fraction analysis. **Flow Measurement and Instrumentation**, v. 14, n. 4–5, p. 161–168, 2003.

GMYTECO, Alexandre Cunningham. **Projeto executivo de um circuito experimental para operar com gases de alto peso molecular**. Trabalho de Conclusão de Curso (Graduação em Engenharia Mecânica). Curitiba: Universidade Tecnológica Federal do Paraná (UTFPR), 2018.

GONZALEZ, Rafael; WOODS, Richad. **Processamento digital de imagens**. São Paulo: Pearson, 2009. v. 3

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning**. Cambridge, MA: MIT Press, 2015.

GOPAL, M.; JEPSON, W. P. Development of digital image analysis techniques for the study of velocity and void profiles in slug flow. **International Journal of Multiphase Flow**, v. 23, n. 5, p. 945–965, mar. 1997.

HE, Kaiming *et al.* **Mask R-CNN**. [S.l.: S.n.]. Disponível em: <<http://arxiv.org/abs/1703.06870>>.

HEBB, D. O. **The organization of behavior: a neuropsychological theory**. New York: Wiley, 1949.

HOFFMANN, João Eduardo *et al.* Real-time adaptive object detection and tracking for autonomous vehicles. **IEEE Transactions on Intelligent Vehicles**, v. 6, n. 3, p. 450–459, 1 set. 2021.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. **ACM Computing Surveys**, v. 31, n. 3, p. 264–323, 1999.

JI, Xu; HENRIQUES, João F.; VEDALDI, Andrea. **Invariant information clustering for unsupervised image classification and segmentation**. [S.l.: S.n.]. Disponível em: <<http://arxiv.org/abs/1807.06653>>.

KADRI, U.; MUDDE, R. F.; OLIEMANS, R. V. A. Influence of the operation pressure on slug length in near horizontal gas-liquid pipe flow. **International Journal of Multiphase Flow**, v. 36, n. 5, p. 423–431, maio 2010.

KALMAN, R. A new approach to linear filtering and prediction problems. **Journal of Basic Engineering**, v. 82, p. 35–45, 1960.

KASS, Michael; WITKIN, Andrew; TERZOPOULOS, Demetri. Snakes: active contour models. **International Journal of Computer Vision**, v. 1, p. 321–331, 1988.

KIRILLOV, Alexander *et al.* Segment Anything. *In*: IEEE, 5 abr. 2023. Disponível em: <<http://arxiv.org/abs/2304.02643>>

KONG, Ran *et al.* A robust image analysis technique for the study of horizontal air-water plug flow. **Experimental Thermal and Fluid Science**, v. 102, p. 245–260, 1 abr. 2019.

KUHN, Harold W. The Hungarian method for the assignment problem. **Naval Research Logistics Quarterly**, v. 2, n. 1–2, p. 83–97, mar. 1955.

LECUN, Y. *et al.* Backpropagation applied to handwritten zip code recognition. **Neural Computation**, v. 1, p. 541–551, 1989.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 27 maio 2015.

LIU, Huishu *et al.* Slug flow hydrodynamics modeling for gas-liquid two-phase flow in a pipe. **Energies**, v. 15, n. 2, p. 533, 1 jan. 2022.

LUO, Wenhan *et al.* Multiple object tracking: a literature review. **Artificial Intelligence**, v. 293, p. 103448, 11 fev. 2022.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. *In*: Berkeley: University of California Press, 1967.

MATHERON, G. **Éléments pour une théorie des milieux poreux**. Paris: Masson et Cie, 1967.

MATHUR, Garima; SOMWANSHI, Devendra; BUNDELE, Mahesh. Intelligent Video Surveillance based on Object Tracking. *In: Jaipur: IEEE, 2018.*

MAYOR, T. S.; PINTO, A. M. F. R.; CAMPOS, J. B. L. M. An image analysis technique for the study of gas-liquid slug flow along vertical pipes - associated uncertainty. **Flow Measurement and Instrumentation**, v. 18, n. 3–4, p. 139–147, jun. 2007.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mothemnticnl Biology**, v. 5, n. 2, p. 115–133, 1943.

METZKER, Leonardo Fadel *et al.* Computer vision techniques for Taylor bubble detection and velocity measurement using YOLO v8 and optical flow. **Flow Measurement and Instrumentation**, v. 104, p. 102885, 1 ago. 2025.

MILLETARI, Fausto; NAVAB, Nassir; AHMADI, Seyed-Ahmad. V-Net: fully convolutional neural networks for volumetric medical image segmentation. *In: IEEE, 15 jun. 2016. Disponível em: <<http://arxiv.org/abs/1606.04797>>*

MINKY, Marvin; PAPER, Seymour. **Perceptrons: an introduction to computational geometry**. Cambridge, MA: MIT Press, 1969.

MOHMED, Abdalellah O. *et al.* Measurements of translational slug velocity and slug length using an image processing technique. **Flow Measurement and Instrumentation**, v. 50, p. 112–120, jun. 2016.

NICKLIN, D. J. Two-phase bubble flow. **Chemical Engineering Science**, v. 17, p. 693–702, 1962.

OKTAY, Ozan *et al.* **Attention U-Net: learning where to look for the pancreas**. *[S.l.: S.n.]*. Disponível em: <<http://arxiv.org/abs/1804.03999>>.

OLBRICH, Marc *et al.* Deep learning based liquid level extraction from video observations of gas–liquid flows. **International Journal of Multiphase Flow**, v. 157, 1 dez. 2022.

O'SHEA, Keiron; NASH, Ryan. **An introduction to convolutional neural networks**. *[S.l.: S.n.]*. Disponível em: <<http://arxiv.org/abs/1511.08458>>.

OTSU, Nobuyuki. A threshold selection method from gray-level histograms. **IEEE Transactions on Systems, Man, and Cybernetics**, n. 1, p. 62–66, 1979.

PERKINS, Sean J. Elongated bubble centring and high-viscosity liquids in horizontal gas-liquid slug flow: Empirical analyses and novel theory. **International Journal of Multiphase Flow**, p. 105430, set. 2025.

PINTO, Walisson Chaves Ferreir *et al.* Bubble detection with semantic segmentation for multiphase flow particle image velocimetry. *In: Salvador, BA: SBIC, 2023. Disponível em: <https://sbia.org.br/wp-content/uploads/2023/10/pdf/CBIC_2023_paper030.pdf>. Acesso em: 2 mar. 2026*

PIPA, Daniel R. *et al.* Typical bubble shape estimation in two-phase flow using inverse problem techniques. **Flow Measurement and Instrumentation**, v. 40, p. 64–73, 1 dez. 2014.

POLONSKY, S.; BARNEA, D.; SHEMER, L. Averaged and time-dependent characteristics of the motion of an elongated bubble in a vertical pipe. **International Journal of Multiphase Flow**, v. 25, p. 795–812, 1999.

QUINTINO, André Mendes; DA FONSECA JUNIOR, Roberto; RODRIGUEZ, Oscar Mauricio Hernandez. Experimental study of liquid/dense-gas pipe flow. **Geoenergy Science and Engineering**, v. 230, 1 nov. 2023.

RADFORD, Alec *et al.* learning transferable visual models from natural language supervision. *In*: PMLR, 2021. Disponível em: <<https://github.com/OpenAI/CLIP>>

RAVI, Nikhila *et al.* **SAM 2: Segment Anything in images and videos**. [S.l.: S.n.]. Disponível em: <<http://arxiv.org/abs/2408.00714>>.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-Net: convolutional networks for biomedical image segmentation. *In*: Cham: Springer International Publishing, 18 maio 2015a. Disponível em: <<http://arxiv.org/abs/1505.04597>>

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-Net: convolutional networks for biomedical image segmentation. *In*: Cham: Springer, 18 maio 2015b. Disponível em: <<http://arxiv.org/abs/1505.04597>>

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological Review**, v. 65, n. 6, p. 368–408, 1958.

RUMELHART, David E.; HINTON, Geoffrey E.; WILLIAMS, Ronald J. Learning representations by backpropagating errors. **Nature**, v. 323, n. 6088, p. 533–536, 1986.

SAUVOLA, J.; PIETIKAK, M. Adaptive document image binarization. **Pattern Recognition**, v. 33, p. 225–236, 2000.

SCHACHTER, Bruce J.; DAMS, Larry S.; ROSENFELD, Azriel. Some experiments in image segmentation by clustering of local feature values. **Pattern Recognition**, v. 1, n. 1, p. 19–28, 1978.

SCHRODTT, Verle N.; SAUNDERST, A. M. Interactive image processing in research. **Computers & Chemical Engineering**, v. 5, n. 4, p. 299–305, 1981.

SHARAF, Amer. A review of image segmentation strategies from classical methods to deep learning. *In*: Saint Petersburg: IEEE, 2024.

SHOHAM, Ovadia. **Mechanistic modeling og gas-liquid two-phase flow in pipes**. Richardson, Texas: Society of Petroleum Engineers (SPE), 2006a.

SHOHAM, Ovadia. **Mechanistic modeling of gas-liquid two-phase flow in pipes.** [S.l.: S.n.].

SILVERMAN, Judith F.; COOPER, David B. Bayesian clustering for unsupervised estimation of surface and texture models. **IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE**, v. 10, n. 4, p. 482–495, 1988.

SOBEL, Irwin. **A 3×3 isotropic gradient operator for image processing.** [S.l.: S.n.]. Disponível em: <<https://www.researchgate.net/publication/285159837>>.

SUTSKEVER, Ilya; VINYALS, Oriol; LE, Quoc V. Sequence to sequence learning with neural networks. *In*: Curran Associates, 14 dez. 2014. Disponível em: <<http://arxiv.org/abs/1409.3215>>

TAITEL, Yehuda; BARNEA, Dvora. Two-phase slug flow. *In*: **Advances in Heat Transfer**. San Diego: Academic Press, 1990. v. 20 p. 83–132.

TAITEL, Yemada; DUKLER, A. E. A model for predicting flow regime transitions in horizontal and near horizontal gas-liquid flow. **AIChE Journal**, v. 22, n. 1, p. 47–55, 1976.

VERSLUIS, Michel. High-speed imaging in fluids. **Experimental Fluids**, v. 54, n. 2, 2012.

VICENT, Luc; SOILLE, Pierre. Watersheds in digital spaces an efficient algorithm based on immersion simulations. v. 13, n. 6, p. 583–598, 1991.

WALLIS, Graham B. **One-dimensional two-phase flow**. New York: McGraw-Hill, 1969.

WANG, Qianwen *et al.* Bubble recognizing and tracking in a plate heat exchanger by using image processing and convolutional neural network. **International Journal of Multiphase Flow**, v. 138, p. 103593, 1 maio 2021.

WANG, Yuxuan *et al.* Tacotron: Towards End-to-End Speech Synthesis. *In*: ISCA, 6 abr. 2017. Disponível em: <<http://arxiv.org/abs/1703.10135>>

WIDYATAMA, Arif *et al.* The development of image processing technique to study the interfacial behavior of air-water slug two-phase flow in horizontal pipes. **Flow Measurement and Instrumentation**, v. 59, p. 168–180, 1 mar. 2018.

WOJKE, Nicolai; BEWLEY, Alex; PAULUS, Dietrich. Simple online and realtime tracking with a deep association metric. *In*: IEEE, 21 mar. 2017. Disponível em: <<http://arxiv.org/abs/1703.07402>>

WU, Zhenyu; LEAHY, Richard. An optimal graph theoretic approach to data clustering: theory and its application to image segmentation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 15, n. 11, p. 1101–1113, 1993.

YASEEN, Muhammad. **What is YOLOv8: an in-depth exploration of the internal features of the next-generation object detector.** [S.l.: S.n.]. Disponível em: <<http://arxiv.org/abs/2408.15857>>.

YILMAZ, Alper; JAVED, Omar; SHAH, Mubarak. Object tracking: a survey. **ACM Computing Surveys**, v. 38, n. 4, 25 dez. 2006.

ZHOU, Zongwei *et al.* UNet++: a nested u-net architecture for medical image segmentation. *In*: Cham: Springer, 18 jul. 2018. Disponível em: <<http://arxiv.org/abs/1807.10165>>

ZOU, Xueyan *et al.* **Segment everything everywhere all at once.** [S.l.: S.n.]. Disponível em: <<http://arxiv.org/abs/2304.06718>>.

APÊNDICE A - Tabelas complementares das baterias experimentais

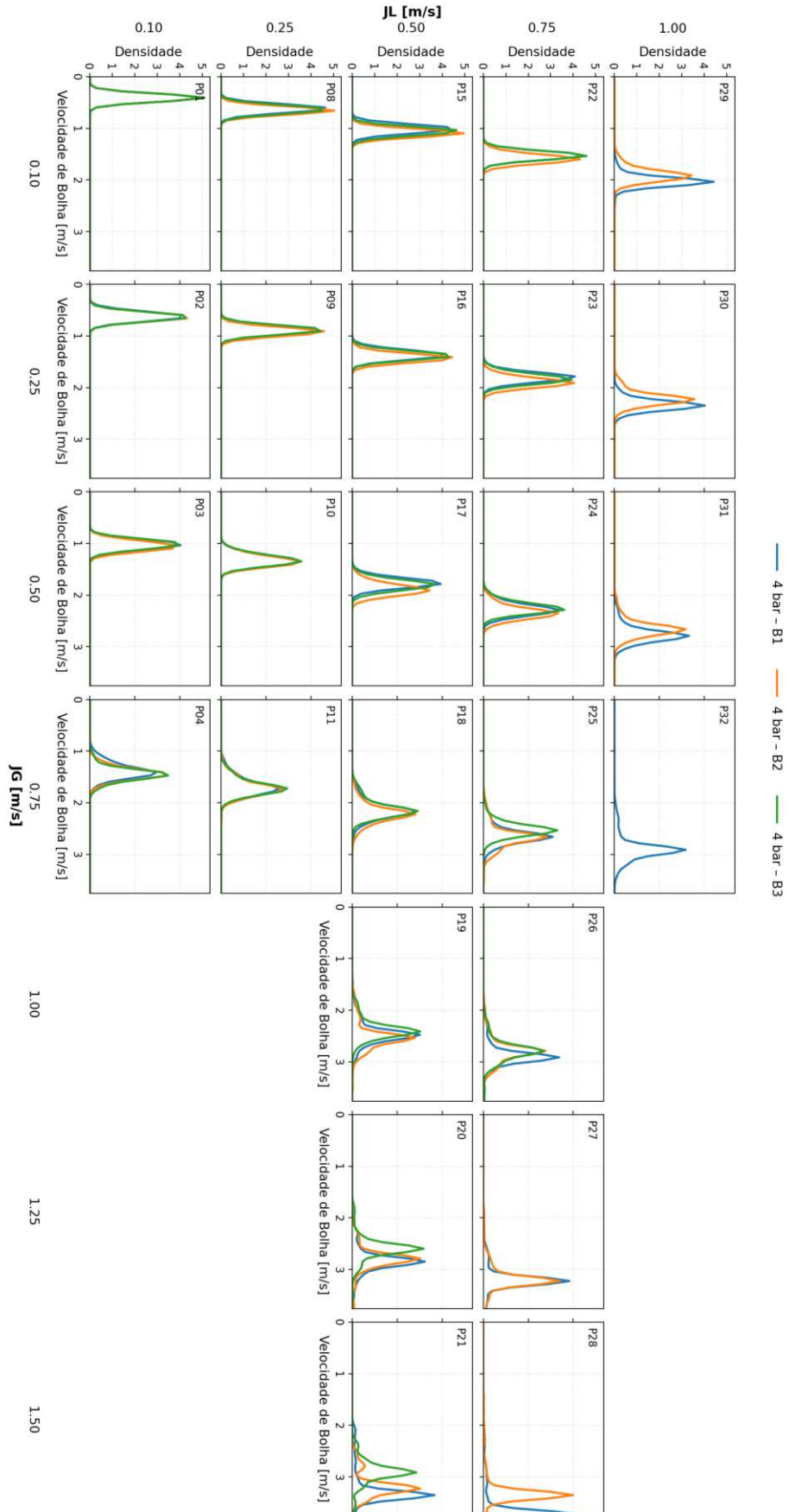
Tabela 5 – Baterias experimentais a 4 bar

Bateria	Padrão	B1				B2				B3			
		JL [m/s]	JG [m/s]	P [bar]	T [°C]	JL [m/s]	JG [m/s]	P [bar]	T [°C]	JL [m/s]	JG [m/s]	P [bar]	T [°C]
P01	Golfada	0,10	0,11	3,94	32,4	0,10	0,11	3,32	29,2	0,10	0,11	3,55	35,7
P02	Golfada	0,11	0,25	3,99	33,4	0,10	0,25	3,33	30,1	0,11	0,25	3,43	34,6
P03	Golfada	0,11	0,50	3,60	31,4	0,11	0,50	3,39	30,1	0,10	0,50	3,61	35,1
P04	Pseudo-Golfada	0,11	0,75	3,65	32,4	0,10	0,75	3,43	30,2	0,10	0,75	3,53	33,9
P05	Estratificado	0,11	1,01	3,69	32,4	0,10	1,01	3,45	30,3	-	-	-	-
P06	Estratificado	0,11	1,26	3,74	32,5	0,10	1,25	3,50	30,3	-	-	-	-
P07	Estratificado	0,10	1,50	3,78	32,5	0,10	1,50	3,55	30,5	-	-	-	-
P08	Golfada	0,25	0,11	3,56	33,4	0,26	0,10	3,19	31,4	0,26	0,10	3,58	34,7
P09	Golfada	0,25	0,26	3,32	32,8	0,26	0,26	3,22	32,3	0,26	0,26	3,60	34,8
P10	Golfada	0,25	0,51	3,51	33,2	0,25	0,49	3,26	32,4	0,25	0,50	3,54	34,6
P11	Pseudo-Golfada	0,25	0,76	3,56	33,4	0,25	0,75	3,32	32,4	0,25	0,75	3,58	34,6
P12	Estratificado	0,25	0,99	3,62	33,4	0,25	1,01	3,39	32,5	-	-	-	-
P13	Estratificado	0,25	1,25	3,65	33,6	0,25	1,25	3,43	32,5	-	-	-	-
P14	Estratificado	0,25	1,50	3,70	33,9	0,25	1,48	3,48	32,5	-	-	-	-
P15	Golfada	0,50	0,10	3,55	34,6	0,51	0,10	3,23	30,2	0,51	0,10	3,65	34,6
P16	Golfada	0,50	0,25	3,57	34,6	0,50	0,26	3,25	31,2	0,50	0,26	3,57	34,6
P17	Golfada	0,50	0,51	3,62	34,8	0,50	0,51	3,31	31,3	0,50	0,50	3,76	35,2
P18	Golfada	0,50	0,76	3,49	35,6	0,50	0,75	3,35	31,4	0,50	0,75	3,67	34,6
P19	Golfada	0,50	1,04	3,57	35,7	0,51	0,99	3,41	31,4	0,51	0,99	3,85	35,7
P20	Golfada	0,50	1,26	3,63	34,6	0,50	1,25	3,33	32,4	-	-	-	-
P21	Golfada	0,50	1,52	3,61	35,7	0,50	1,48	3,43	32,4	-	-	-	-
P22	Golfada	0,75	0,11	3,60	36,8	0,75	0,11	3,29	30,5	0,75	0,11	3,73	35,7
P23	Golfada	0,75	0,25	3,62	36,8	0,75	0,25	3,32	31,2	0,75	0,25	3,67	35,5
P24	Golfada	0,75	0,50	3,58	34,5	0,75	0,50	3,32	31,4	0,75	0,50	3,73	35,7
P25	Golfada	0,76	0,75	3,54	35,3	0,75	0,75	3,34	32,3	0,75	0,74	3,91	36,6
P26	Golfada	0,75	0,97	3,60	34,6	0,75	0,98	4,01	35,7	-	-	-	-
P27	Golfada	0,75	1,25	3,61	35,7	0,75	1,28	3,93	35,6	-	-	-	-
P28	Golfada	0,75	1,52	3,60	35,7	0,75	1,53	4,03	35,7	-	-	-	-
P29	Pseudo-Golfada	1,01	0,10	3,28	35,7	0,99	0,11	3,81	36,8	-	-	-	-
P30	Golfada	1,01	0,26	3,29	36,8	1,01	0,25	3,86	36,5	-	-	-	-
P31	Golfada	1,00	0,50	3,31	36,8	1,00	0,50	3,95	36,8	-	-	-	-
P32	Golfada	1,00	0,74	3,31	37,9	-	-	-	-	-	-	-	-

Fonte: Autoria própria.

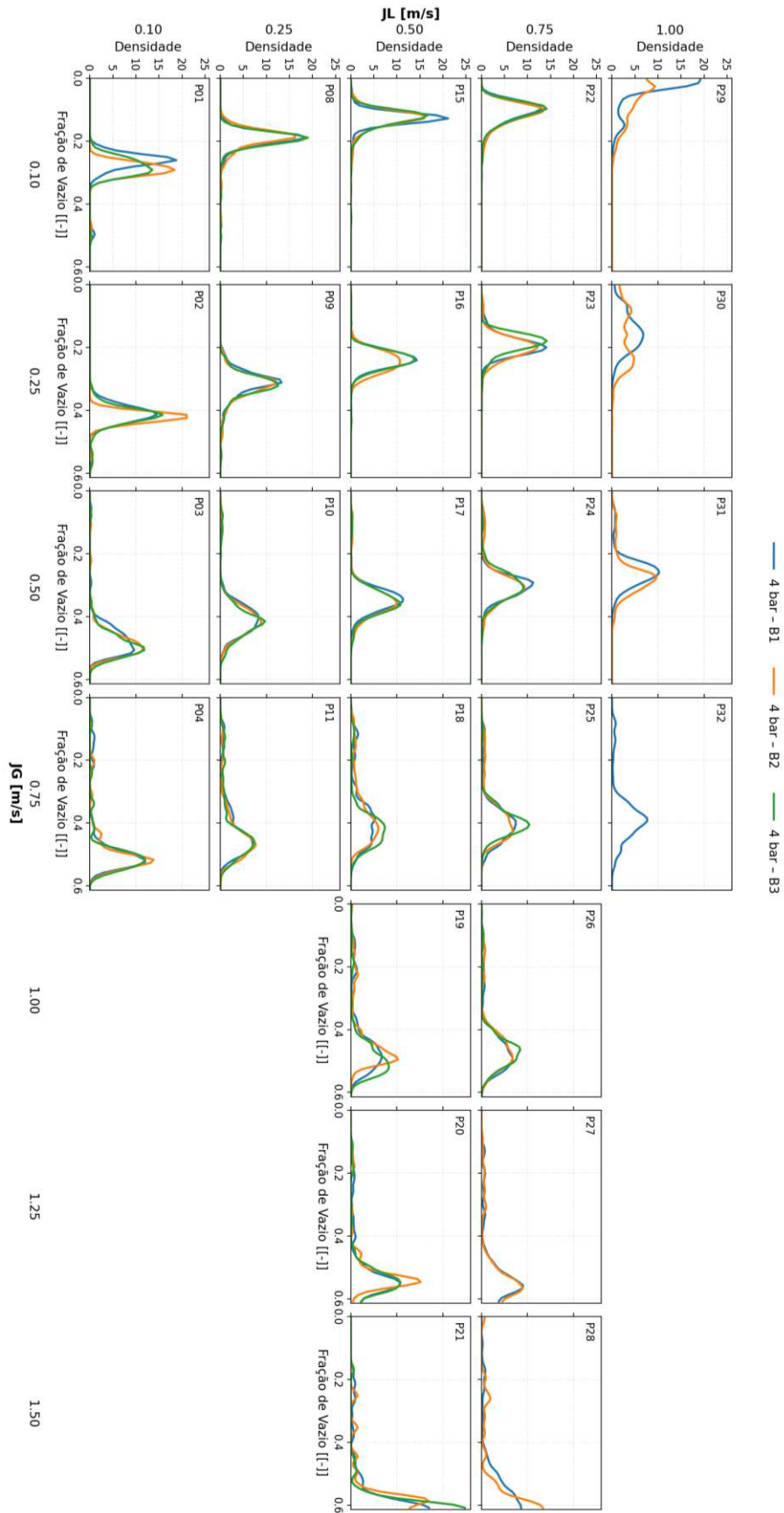
APÊNDICE B - Histogramas em matriz dos parâmetros do escoamento (4 bar)

Figura 33 – Histograma da velocidade de bolha



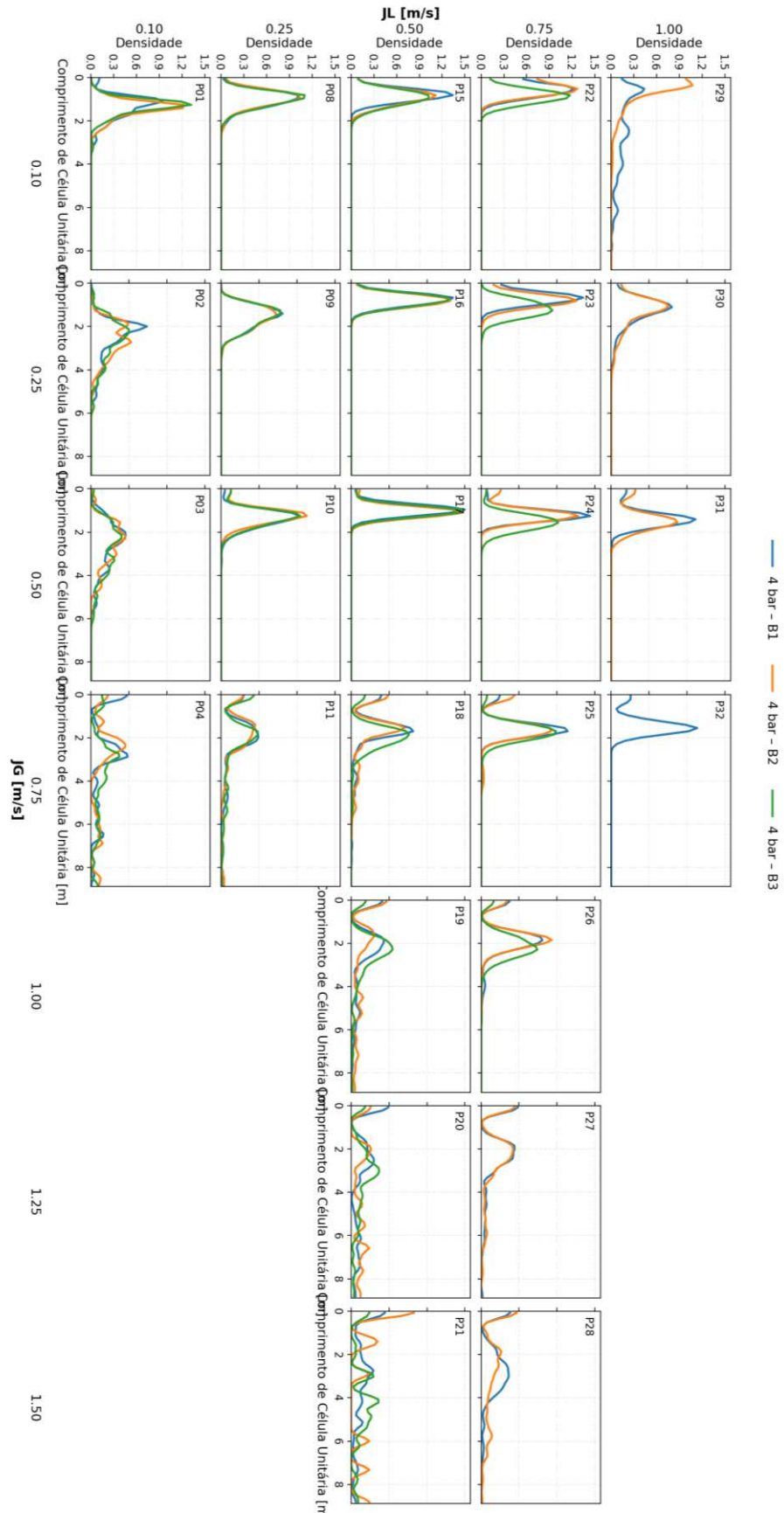
Fonte: Autoria própria

Figura 34 – Histograma da fração de vazio



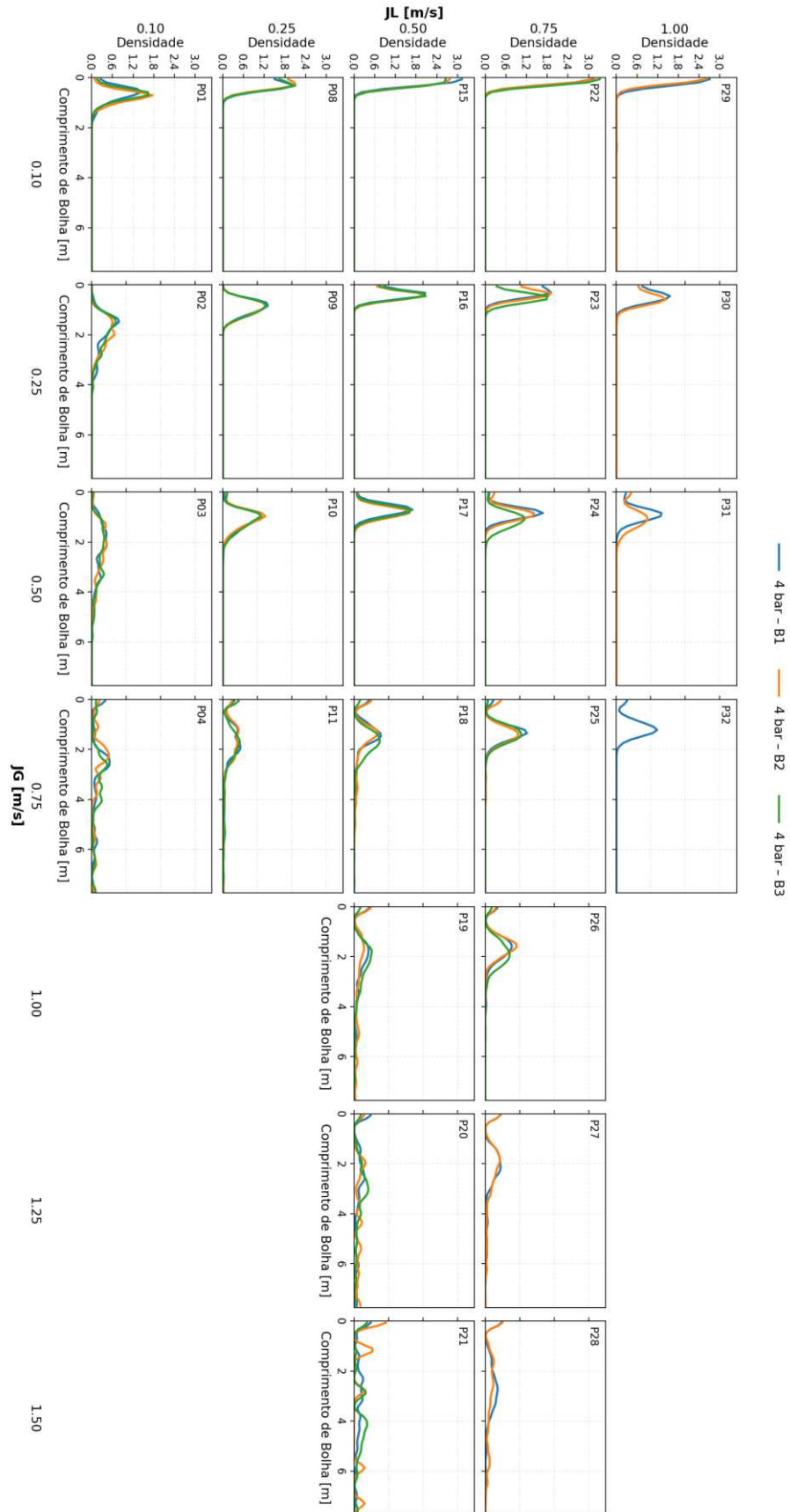
Fonte: Autoria própria

Figura 35 – Histograma do comprimento de célula unitária



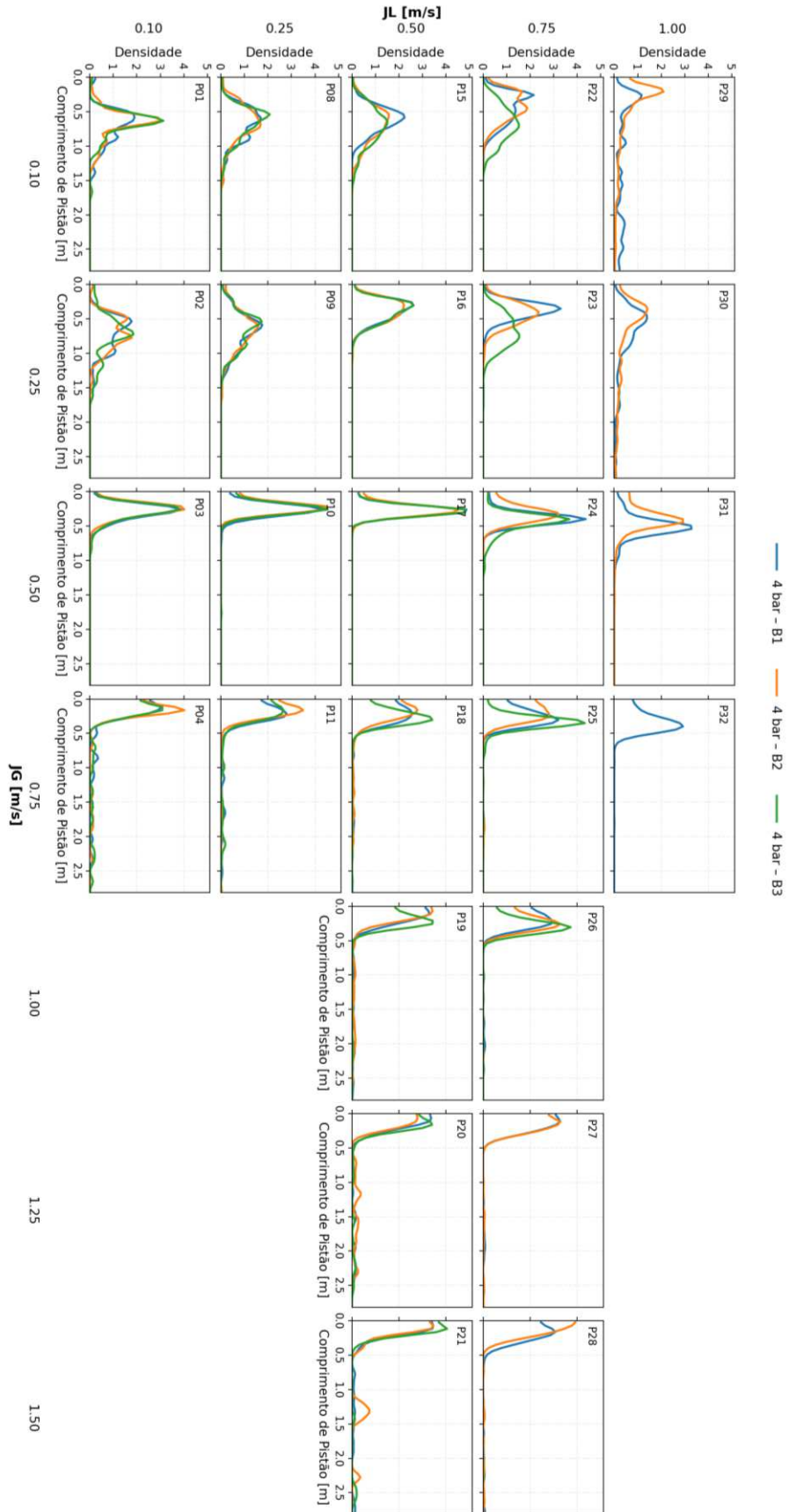
Fonte: Autoria própria

Figura 36 – Histograma de comprimento de bolha



Fonte: Autoria própria

Figura 37 – Histograma de comprimento de pistão de líquido



Fonte: Autoria própria