



Universidade Estadual de Campinas
Faculdade de Engenharia Química



Leonardo Fadel Metzker

Visão Computacional Aplicada à Detecção e à Caracterização de Bolhas de Taylor em Meio Infinito

Campinas - SP
2024

LEONARDO FADEL METZKER

VISÃO COMPUTACIONAL APLICADA À
DETECÇÃO E À CARACTERIZAÇÃO DE
BOLHAS DE TAYLOR EM MEIO INFINITO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Química da Universidade Estadual de Campinas, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Química.

Orientadora: Profa. Dra. Ana Maria Frattini Fileti

Coorientador: Prof. Dr. Maurício de Melo Freire Figueiredo

Este arquivo digital corresponde à versão final da dissertação defendida por Leonardo Fadel Metzker e orientada pela Profa. Dra. Ana Maria Frattini Fileti.

Campinas - SP

2024

Ficha catalográfica
Universidade Estadual de Campinas (UNICAMP)
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

M569v Metzker, Leonardo Fadel, 1998-
Visão computacional aplicada à detecção e à caracterização de bolhas de Taylor em meio infinito / Leonardo Fadel Metzker. – Campinas, SP : [s.n.], 2024.

Orientador(es): Ana Maria Frattini Fileti.
Coorientador(es): Maurício de Melo Freire Figueiredo.
Dissertação (mestrado) – Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia Química.

1. Visão por computador. 2. Fluxo óptico. 3. Medidores de fluxo. 4. Classificação de imagem. I. Fileti, Ana Maria Frattini, 1965-. II. Figueiredo, Maurício de Melo Freire, 1985-. III. Universidade Estadual de Campinas (UNICAMP). Faculdade de Engenharia Química. IV. Título.

Informações complementares

Título em outro idioma: Computer vision applied to the detection and characterization of Taylor bubbles in an infinite medium

Palavras-chave em inglês:

Computer vision

Optical flow

Flow meters

Image classification

Área de concentração: Engenharia Química

Titulação: Mestre em Engenharia Química

Banca examinadora:

Ana Maria Frattini Fileti [Orientador]

Ricardo Augusto Mazza

Silvio Jamil Ferzoli Guimarães

Data de defesa: 21-11-2024

Programa de Pós-Graduação: Engenharia Química

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0009-0007-1716-0683>

- Currículo Lattes do autor: <https://lattes.cnpq.br/7423601297369212>

Folha de Aprovação da Defesa de Dissertação de Mestrado defendida por **Leonardo Fadel Metzker, 178440** e aprovada em 21 de novembro de 2024 pela Comissão Examinadora constituída pelos doutores:

Profa. Dra. Ana Maria Frattini Fileti
(Presidente) (Orientadora)
FEQ/ UNICAMP

Prof. Dr. Ricardo Augusto Mazza
FEM/ UNICAMP

Dr. Silvio Jamil Ferzoli Guimarães
PUC - Minas Gerais
Belo Horizonte/ MG

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA / Sistema de Fluxo de Dissertação / Tese e na Secretaria do Programa da Unidade.

AGRADECIMENTOS

A conclusão deste trabalho foi possível graças ao apoio incondicional de pessoas essenciais em minha vida. Aos meus pais, cuja dedicação, amor e ensinamentos moldaram não apenas meu caráter, mas também meu percurso acadêmico e profissional, expresso minha mais profunda gratidão. Aos amigos e familiares, que estiveram ao meu lado, oferecendo palavras de encorajamento e compartilhando momentos de superação e alegria, deixo registrado meu sincero reconhecimento.

Gostaria de expressar minha gratidão à minha orientadora, Ana Maria Frattini Fileti, por sua orientação, paciência e valiosas contribuições ao longo deste trabalho. Ao meu co-orientador, Maurício de Melo Freire Figueiredo, pelo apoio técnico e incentivo em momentos cruciais. E ao meu colega de laboratório, Cáio César Silva Araújo, pelo companheirismo, troca de ideias e suporte durante todo o processo de pesquisa.

Agradeço também ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo suporte financeiro, cuja contribuição foi indispensável para a realização deste trabalho, concedido por meio da bolsa processo nº 131637/2021-5.

RESUMO

Escoamentos bifásicos são cruciais em setores industriais como o petroquímico e o de energia nuclear, exigindo uma compreensão das interações entre as fases. Este estudo investigou o uso de visão computacional para medições não intrusivas em escoamentos líquido-gás translúcidos, com foco na análise da velocidade de bolhas de Taylor. Foram utilizadas técnicas de detecção de objetos por redes neurais convolucionais (YOLOv8) integradas ao método de fluxo óptico Lucas-Kanade. Um algoritmo foi desenvolvido para processar filmagens obtidas no *Flow and Research Laboratory* (Flow&Rs) do Centro de Estudos de Energia e Petróleo (CEPETRO) da Universidade Estadual de Campinas (Unicamp). O detector apresentou uma precisão média de 88% para a identificação de bolhas e alcançou uma velocidade média de processamento de 137 FPS. As medições de velocidade apresentaram uma diferença inferior a 6% em relação aos valores reportados na literatura e menor que 1% em comparação aos resultados experimentais. A constante de velocidade terminal calculada foi de 0,34, posicionando-se em concordância com a literatura. Os resultados indicam que a abordagem desenvolvida é promissora para estudos de escoamentos bifásicos envolvendo bolhas de Taylor em meio infinito. Este trabalho confirma a hipótese de que a análise de velocidade por técnicas computacionais oferece uma alternativa eficaz às metodologias tradicionais.

ABSTRACT

Biphasic flows are crucial in industrial sectors such as petrochemical and nuclear energy, requiring a deep understanding of phase interactions. This study investigated the use of computer vision for non-intrusive measurements in translucent liquid-gas flows, focusing on the velocity analysis of Taylor bubbles. Object detection techniques based on convolutional neural networks (YOLOv8) were integrated with the Lucas-Kanade optical flow method. An algorithm was developed to process videos obtained at the Flow and Research Laboratory (Flow&Rs) of the Center for Energy and Petroleum Studies (CE-PETRO) at the University of Campinas (Unicamp). The detector achieved an average accuracy of 88% for bubble identification and an average processing speed of 137 FPS. Velocity measurements showed a deviation of less than 6% compared to values reported in the literature and less than 1% compared to experimental results. The calculated terminal velocity constant was 0.34, aligning with the literature. The results indicate that the developed approach is promising for studies of biphasic flows involving Taylor bubbles in infinite media. This work confirms the hypothesis that velocity analysis through computational techniques offers an effective alternative to traditional methodologies.

SUMÁRIO

1	Introdução	10
2	Objetivo	10
3	Fundamentação Teórica	12
3.1	Escoamento Bifásico Vertical	12
3.1.1	Escoamento Pistonado	13
3.2	Visão Computacional	16
3.2.1	Leitura de Imagens	18
3.2.2	Redes Neurais Convolucionais	18
3.2.3	Modelos de Detecção de Objetos	20
3.2.4	Algoritmo de Detecção de Objetos YOLO	24
3.2.5	Método de Fluxo Óptico	28
3.3	Métodos Estatísticos	35
3.3.1	Média e Desvio Padrão	35
3.3.2	Distribuição em Violino	37
3.4	Conclusão	38
4	Revisão Bibliográfica	39
4.1	Métodos de Classificação e Detecção	39
4.2	Método de Fluxo Ótico	40
4.3	Integração de Detector com Método de Fluxo Óptico	41
4.4	Conclusão	42
5	Metodologia	44
5.1	Montagem Experimental	44
5.2	Procedimento Experimental	44
5.3	Detector de Objetos	47
5.4	Algoritmo de Rastreamento por Fluxo Óptico	48
5.5	Tratamento Matemático dos Resultados	51
5.6	Funcionamento Global do Algoritmo	53
5.7	Descrição do Hardware e Versões de Software	55
6	Resultados	56
6.1	Treinamento do Detector de Objetos	56
6.2	Comprimento das Bolhas de Taylor	57
6.3	Velocidade e Constante de Velocidade Terminal das Bolhas de Taylor	61
6.3.1	Aplicação Delimitada do Método de Fluxo Óptico	61

6.3.2	Avaliação da Dispersão dos Dados de Velocidade e Constante Terminal	63
6.3.3	Comparação com a Literatura	68
6.3.4	Comparação com a Literatura	68
6.3.5	Velocidade de Processamento do Algoritmo	70
6.3.6	Constante de Velocidade Terminal	72
7	Conclusão	74
7.1	Trabalhos Futuros	76

1 Introdução

A presença de fluxos multifásicos em diversos processos industriais é amplamente investigada devido à necessidade de compreender o comportamento desses escoamentos complexos em operações fabris. Entre os fluxos multifásicos, os escoamentos bifásicos, compostos por duas fases fluindo simultaneamente, são particularmente relevantes em setores como as indústrias química, alimentícia, petrolífera e nuclear (XIAOWEI, 2009).

A interação entre as fases em um escoamento multifásico é influenciada por variáveis como propriedades físicas, concentração, pressão, regime de fluxo, geometria do sistema de tubulação e velocidade. Essas interações resultam em diferentes padrões de escoamento, que dependem das condições do processo (RAJAN et al., 1993).

No contexto de fluxos bifásicos líquido-gás, um padrão frequentemente observado é o escoamento pistonado em tubulações verticais, caracterizado pela alternância de pistões de líquido e bolhas alongadas, conhecidas como bolhas de Taylor. A análise do comportamento dessas bolhas é essencial para uma compreensão abrangente do escoamento pistonado, fornecendo subsídios para otimizar processos industriais (FALCONE et al., 2009).

A caracterização de fluxos bifásicos apresenta desafios consideráveis, principalmente devido à necessidade de técnicas de medição não intrusivas que não alterem as características do escoamento. A velocidade das fases é uma variável crucial, pois permite identificar padrões de escoamento e avaliar a estabilidade e o desempenho dos processos (AMERAN, 2016). Métodos tradicionais de medição muitas vezes são invasivos e podem aumentar o consumo energético (FALCONE et al., 2009).

Este trabalho propõe uma abordagem baseada na análise de imagens, utilizando visão computacional para detectar, rastrear e medir a velocidade de bolhas de Taylor. Foram empregadas técnicas avançadas, como o algoritmo YOLO para detecção de objetos e o método de fluxo óptico para análise de movimento. Essas metodologias são não intrusivas e oferecem um potencial significativo para aplicação em estudos de escoamentos bifásicos translúcidos (MAYOR et al., 2008).

A hipótese que orienta este estudo é que a análise da velocidade de bolhas em escoamentos bifásicos por técnicas computacionais oferece uma alternativa viável e precisa em comparação aos métodos tradicionais. A validação dessa hipótese é realizada por meio da comparação com dados da literatura e resultados experimentais.

2 Objetivo

Este estudo tem como objetivo geral investigar o uso de técnicas de visão computacional, especificamente o algoritmo YOLOv8 e o método de fluxo óptico Lucas-Kanade,

para a análise de escoamentos líquido-gás contendo bolhas de Taylor em meio infinito. Os objetivos específicos incluem:

1. Avaliar a precisão do algoritmo YOLOv8 na detecção e rastreamento de bolhas de Taylor;
2. Verificar a eficácia do método de fluxo óptico na medição da velocidade das bolhas;
3. Determinar os comprimentos característicos das bolhas de Taylor;
4. Examinar a viabilidade de aplicação dessas técnicas para medições em tempo real.

3 Fundamentação Teórica

Nesta seção, são apresentados os conceitos teóricos fundamentais para o desenvolvimento deste estudo. São abordados os escoamentos bifásicos verticais, com ênfase nas características dos escoamentos contendo bolhas de Taylor e suas respectivas modelagens. Também são discutidas as ferramentas de visão computacional empregadas, como o algoritmo YOLO para detecção de objetos e o método de fluxo óptico para a medição da velocidade das bolhas. Além disso, são detalhados os métodos estatísticos utilizados na validação dos resultados experimentais, incluindo análise de dispersão, médias, medianas e desvio padrão. Esta seção proporciona a base teórica necessária para a compreensão e execução da pesquisa, facilitando a interpretação dos dados e a validação das técnicas de visão computacional aplicadas.

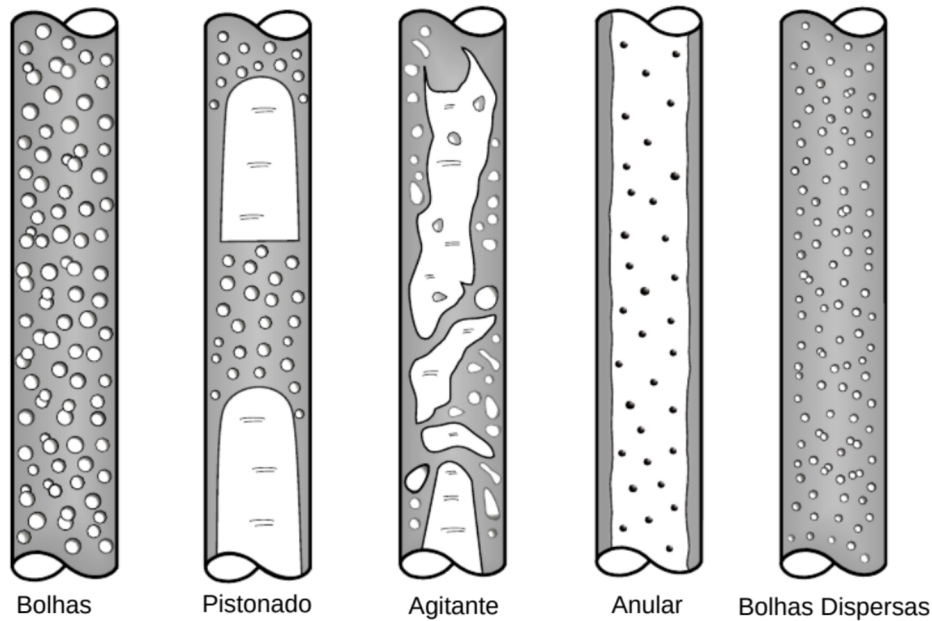
3.1 Escoamento Bifásico Vertical

O escoamento em tubulações verticais de sistemas bifásicos apresenta diferentes distribuições topológicas, comumente denominadas padrões de escoamento. As classificações desses padrões variam conforme diferentes autores. Para este trabalho, serão adotadas as classificações apresentadas por Taitel et al. (1980). A seguir, são apresentados os padrões de escoamento definidos pelos autores e ilustrados na Figura 1, adaptada de (SHOHAM, 2006):

1. **Bolhas:** Bolhas de tamanho mediano que se distribuem de forma homogênea ao longo da tubulação, apresentando uma velocidade superficial baixa, similar à do líquido.
2. **Pistonado:** Neste padrão, formam-se bolhas de Taylor, que são bolhas alongadas que ocupam quase toda a seção da tubulação, deixando apenas uma fina camada de líquido entre a bolha e a parede da tubulação. A velocidade superficial do líquido é relativamente baixa, enquanto a do gás varia de baixa a média.
3. **Agitante:** Semelhante ao padrão pistonado, porém com características geométricas mais caóticas, frequentemente apresentando bolhas distorcidas ao longo do escoamento. A velocidade superficial do líquido é de baixa a média, enquanto a do gás varia de média a alta.
4. **Anular:** Neste padrão, ocorre uma rápida passagem de gás, formando uma região central de gás com um filme de líquido entre o gás e a parede da tubulação, com pequenas bolhas de líquido na parte gasosa. A velocidade superficial do gás é relativamente alta.

5. **Bolhas Dispersas:** Semelhante ao padrão de bolhas, mas com bolhas de dimensões significativamente menores. A velocidade superficial do líquido é relativamente alta, arrastando a fase gasosa.

Figura 1: Padrões de escoamento líquido-gasoso vertical.



3.1.1 Escoamento Pistonado

Fernandes et al. (1983) apresenta um modelo para o padrão pistonado quando este atinge um estado totalmente desenvolvido, ou seja, um estado pseudo-estacionário. Nesse modelo, existem duas regiões delimitadoras para cada passagem: uma é a bolha de Taylor, um aglomerado gasoso caracterizado por sua forma de projétil, que ocupa praticamente toda a seção do tubo, com apenas uma fina camada de líquido entre a bolha e a parede da tubulação (KANG et al., 2010); e a região de *slug*, formada pelo líquido e pelas bolhas de gás que são arrastadas pelo filme de líquido entre a bolha de Taylor e a parede da tubulação, criando uma zona de mistura aerada por pequenas bolhas de gás.

Para descrever o modelo, alguns conceitos de estruturas bifásicas devem ser introduzidos:

- **Vazão mássica total (W):** É a soma das vazões mássicas parciais de líquido (W_L) e gás (W_G) que estão sendo alimentadas no processo:

$$W = W_L + W_G \quad (1)$$

- **Vazão volumétrica total (q):** É a soma das vazões volumétricas parciais de líquido (q_L) e gás (q_G) que estão sendo alimentadas no processo:

$$q = q_L + q_G \quad (2)$$

- **Hold-up de líquido (H_L) e fração de vazios (α):** São as frações que correspondem, respectivamente, à porção de líquido contido na fase e à de gás contido na fase. A soma dessas frações em qualquer porção do escoamento deve ser equivalente a 1:

$$H_L + \alpha = 1 \quad (3)$$

- **Velocidade superficial de líquido (v_{SL}) e gás (v_{SG}):** São as velocidades que ocorreriam caso cada uma das fases fluísse sozinha na tubulação, sendo a vazão volumétrica da fase por unidade de área (A_p):

$$v_{SL} = \frac{q_L}{A_p} \quad (4)$$

$$v_{SG} = \frac{q_G}{A_p} \quad (5)$$

- **Velocidade de mistura (v_M):** É a velocidade central do volume, dada pela soma das velocidades superficiais das fases:

$$v_M = v_{SL} + v_{SG} = \frac{q_G + q_L}{A_p} \quad (6)$$

- **Velocidade real das fases líquida (v_L) e gasosa (v_G):** Como cada fase ocupa apenas uma fração do escoamento, a velocidade real de cada uma delas deve ser calculada considerando apenas a porção em que estas estão presentes:

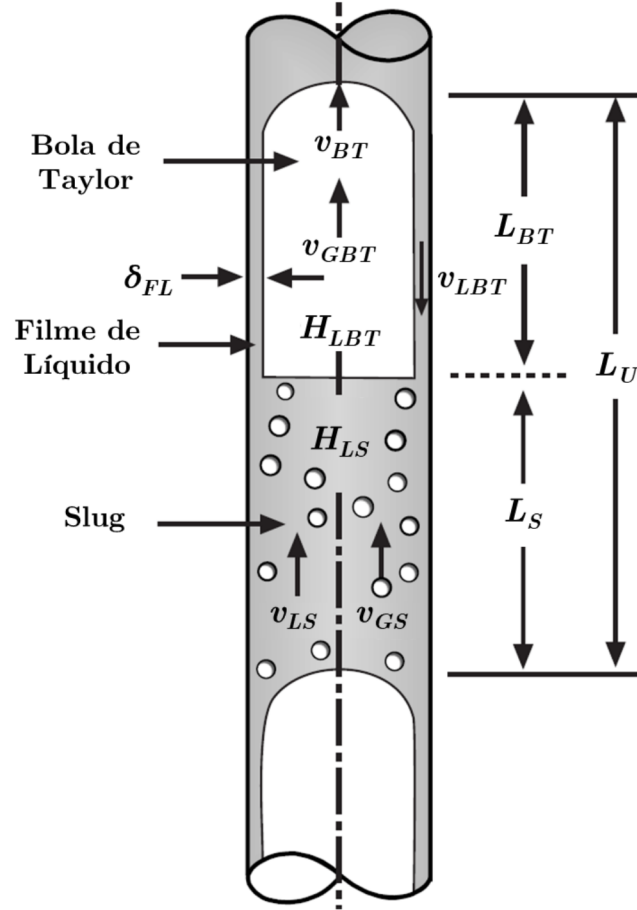
$$v_L = \frac{v_{SL}}{H_L} \quad (7)$$

$$v_G = \frac{v_{SG}}{\alpha} = \frac{v_{SG}}{1 - H_L} \quad (8)$$

Um modelo esquemático para o escoamento pistonado é apresentado na Figura 2, onde está presente a célula unitária, de comprimento L_U , composta pela porção contendo a bolha de Taylor, de comprimento L_{BT} , e o pistão, de comprimento L_S . Na porção da bolha de Taylor, existe um *hold-up* de líquido H_{LBT} , fração correspondente à formação do filme de líquido da ponta até o fim da bolha, com espessura δ_{FL} , e velocidade v_{LBT} .

Além disso, há a presença da bolha em si, que possui uma velocidade translacional v_{BT} e velocidade interna de gás igual a v_{GBT} . Já a fração do pistão apresenta um *hold-up* de líquido H_{LS} , com velocidades de gás e líquido equivalentes a v_{LS} e v_{GS} , respectivamente.

Figura 2: Esquema de um escoamento pistonado desenvolvido, adaptado de (SHOHAM, 2006).



Considerando um balanço de massa global, obtêm-se as equações para as velocidades superficiais de líquido (v_{SLU}) e gás (v_{SGU}) da célula unitária:

$$v_{SLU} = (1 - \beta)v_{LBT}H_{LBT} + (1 - \beta)v_{GS}H_{LS} \quad (9)$$

$$v_{SGU} = \beta v_{GBT}(1 - H_{LBT}) + \beta v_{GS}(1 - H_{LS}) \quad (10)$$

$$\beta = \frac{L_F}{L_U} \quad (11)$$

Outro balanço de massa é realizado considerando as fases líquidas entre as porções da célula unitária, levando em conta as coordenadas a partir de v_{BT} , resultando na equação:

$$(v_{BT} - v_{LS})H_{LS} = (v_{BT} - v_{LBT})H_{LBT} \quad (12)$$

A velocidade de v_{LBT} é apresentada negativa pois, conforme a esquematização, seu sentido é oposto ao de v_{BT} , o que justifica a inclusão do sinal negativo. O mesmo balanço pode ser realizado para a porção gasosa, resultando em:

$$(v_{BT} - v_{GS})(1 - H_{LS}) = (v_{BT} - v_{GBT})(1 - H_{LBT}) \quad (13)$$

A velocidade da bolha de Taylor pode ser expressa pela soma da velocidade máxima do fluido na linha central do escoamento (v_{max}) com a velocidade terminal da bolha (v_t), sendo a velocidade da bolha caso ela escoasse em um meio estagnado:

$$v_{BT} = c_o v_{max} + v_t \quad (14)$$

onde c_o é o parâmetro empírico utilizado para considerar os efeitos de arraste e de forma da bolha.

A avaliação do escoamento de bolhas de Taylor em meio estagnado envolve análises teóricas e experimentais que quantificam a velocidade de ascensão das bolhas em uma tubulação vertical. Uma das primeiras equações para a velocidade terminal (v_t) foi proposta por Dumitrescu (1943), considerando o efeito da gravidade (g), o diâmetro interno da tubulação (d) e uma constante empírica (C_t), conforme:

$$v_t = C_t \cdot \sqrt{gd} \quad (15)$$

Essa constante foi estudada e avaliada por outros autores, resultando em, em alguns casos, valores distintos, como apresentado na Tabela 1.

Tabela 1: Constantes para a velocidade terminal.

Referência	C_t	Método
Dumitrescu (1943)	0,351	Analítico
Davies e Taylor (1950)	0,328	Analítico
Campos e Carvalho (1988)	0,350	Experimental
Laird e Chisholm (1956)	0,328	Experimental
Polonsky et al. (1999)	0,351	Experimental

3.2 Visão Computacional

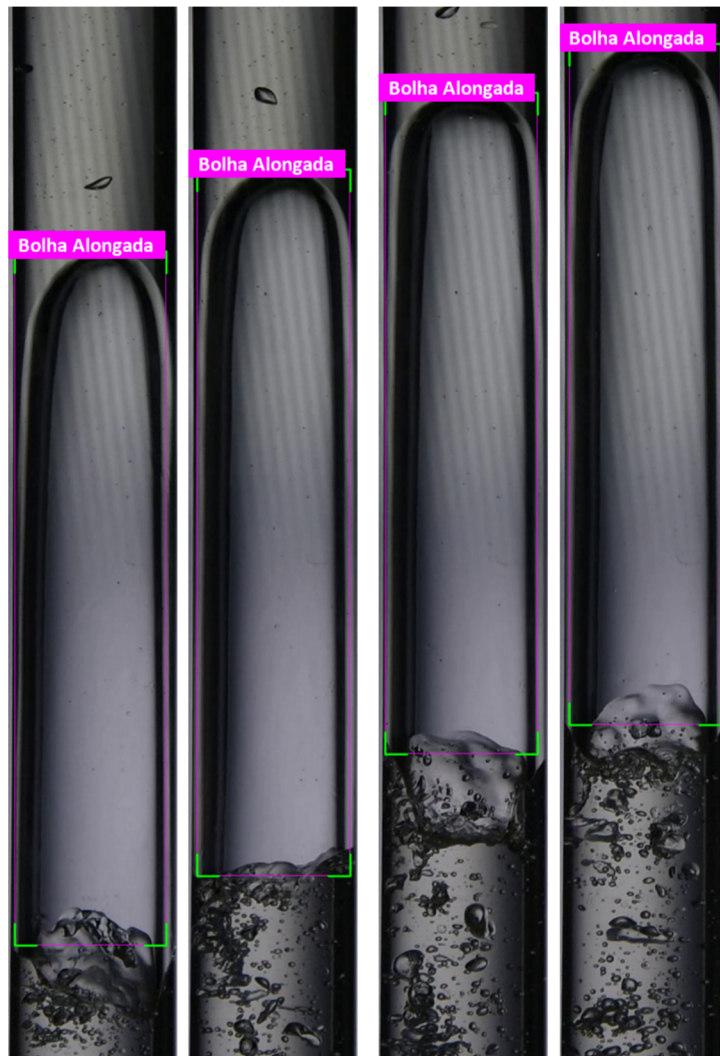
A visão computacional, um ramo da inteligência artificial (IA), capacita os computadores a analisar e compreender imagens, extraindo padrões dos objetos nelas contidos para

interpretar seu conteúdo e, em alguns casos, tomar decisões com base nesses resultados (ANSARI, 2020).

Além da classificação, a visão computacional é aplicada na detecção e rastreamento de objetos em imagens, sendo utilizada em diversos setores, como segurança, imagens médicas, veículos autônomos e reconhecimento facial (KALE et al., 2015).

Essas aplicações frequentemente envolvem a análise de vídeos, compostos por uma sequência de quadros apresentados a uma taxa de quadros por segundo (FPS). O objetivo é detectar objetos nos quadros, determinar sua presença ou ausência ao longo do tempo, rastrear sua trajetória e, em alguns casos, prever seus movimentos futuros. A Figura 3 ilustra um exemplo de detecção e rastreamento de bolhas alongadas.

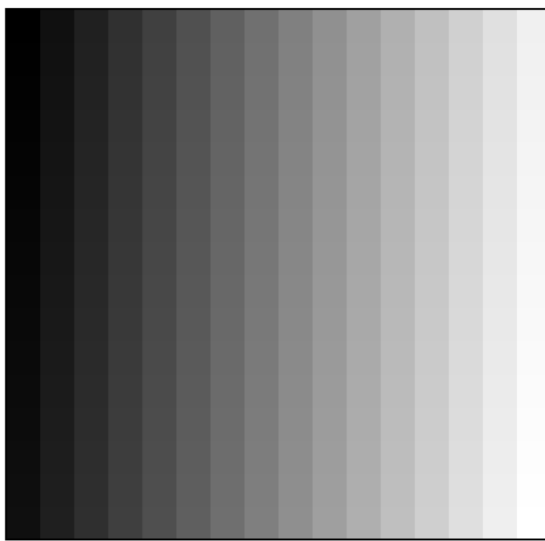
Figura 3: Exemplo de detecção e rastreamento de bolhas alongadas ao longo de um vídeo.



3.2.1 Leitura de Imagens

A leitura de imagens por uma máquina ocorre numericamente, considerando que um pixel é a menor unidade de uma imagem. A resolução da imagem determina o número total de pixels; por exemplo, uma imagem com resolução de 1080x600 possui 648.000 pixels, que é o produto de 1080 linhas de pixels por 600 colunas. A intensidade de cada pixel é usualmente quantizada em uma escala de 8 bits, variando de 0 a 255. A Figura 4 ilustra a quantização de pixels em uma camada de escala de cinza.

Figura 4: Quantização de pixels em uma camada de escala de cinza em 8-bits.



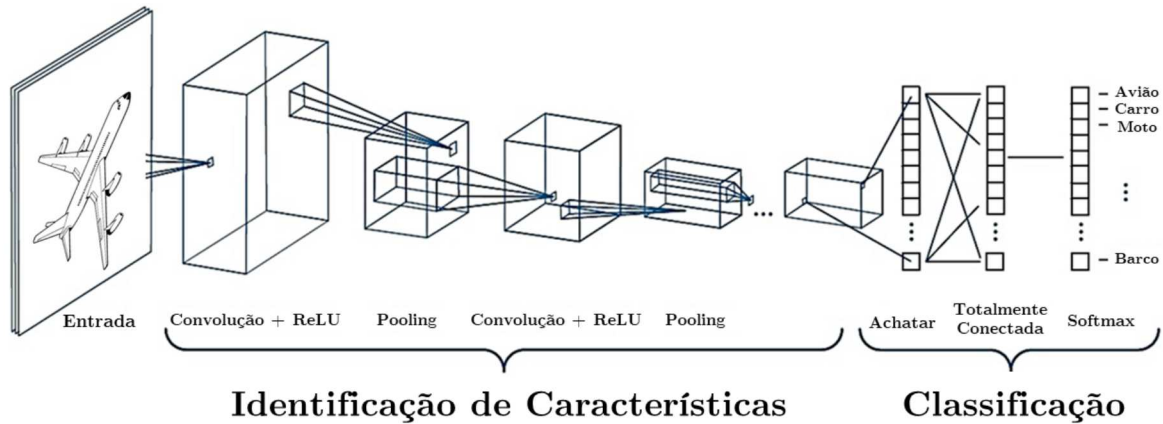
0	16	32	48	64	80	96	112	128	144	160	176	192	208	224	240
1	17	33	49	65	81	97	113	129	145	161	177	193	209	225	241
2	18	34	50	66	82	98	114	130	146	162	178	194	210	226	242
3	19	35	51	67	83	99	115	131	147	163	179	195	211	227	243
4	20	36	52	68	84	100	116	132	148	164	180	196	212	228	244
5	21	37	53	69	85	101	117	133	149	165	181	197	213	229	245
6	22	38	54	70	86	102	118	134	150	166	182	198	214	230	246
7	23	39	55	71	87	103	119	135	151	167	183	199	215	231	247
8	24	40	56	72	88	104	120	136	152	168	184	200	216	232	248
9	25	41	57	73	89	105	121	137	153	169	185	201	217	233	249
10	26	42	58	74	90	106	122	138	154	170	186	202	218	234	250
11	27	43	59	75	91	107	123	139	155	171	187	203	219	235	251
12	28	44	60	76	92	108	124	140	156	172	188	204	220	236	252
13	29	45	61	77	93	109	125	141	157	173	189	205	221	237	253
14	30	46	62	78	94	110	126	142	158	174	190	206	222	238	254
15	31	47	63	79	95	111	127	143	159	175	191	207	223	239	255

No caso de imagens coloridas, um dos modos mais usuais de interpretação é na escala RGB (Red, Green, Blue), composta por três camadas de intensidades diferentes que se sobrepõem para criar a imagem completa.

3.2.2 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (RNCs) são modelos amplamente utilizados em aprendizado profundo, especialmente desenvolvidos para o processamento de imagens. Elas consistem em uma forma especializada de Rede Neural Artificial (RNA), projetada para manipular dados com estrutura espacial correlacionada, em que a proximidade física entre elementos de dados reflete relações de significado. Em outras palavras, as RNCs são particularmente eficazes no reconhecimento de padrões em dados visuais, como imagens, devido à sua capacidade de explorar a estrutura espacial subjacente. Essas redes são amplamente empregadas em aplicações como classificação de imagens e reconhecimento de objetos, possuindo uma estrutura que inclui a extração de características e a etapa de classificação, conforme ilustrado na Figura 7 (KIM et al., 2023).

Figura 5: Exemplo de arquitetura de uma RNC para classificação de meios de transporte por imagem.



A estrutura do modelo baseia-se na análise dos dados de entrada e saída, incorporando várias camadas internas que realizam operações de convolução, ativação, *pooling* e conexões densas, ou totalmente conectadas. As camadas internas processam os dados de forma hierárquica, permitindo a identificação e a extração progressiva de características relevantes, como bordas, texturas e formas.

As conexões densas, ou totalmente conectadas, referem-se a camadas em que cada neurônio está conectado a todos os neurônios da camada subsequente. Essa etapa é fundamental para combinar as características extraídas pelas camadas anteriores e produzir a saída final, como a classificação de uma imagem.

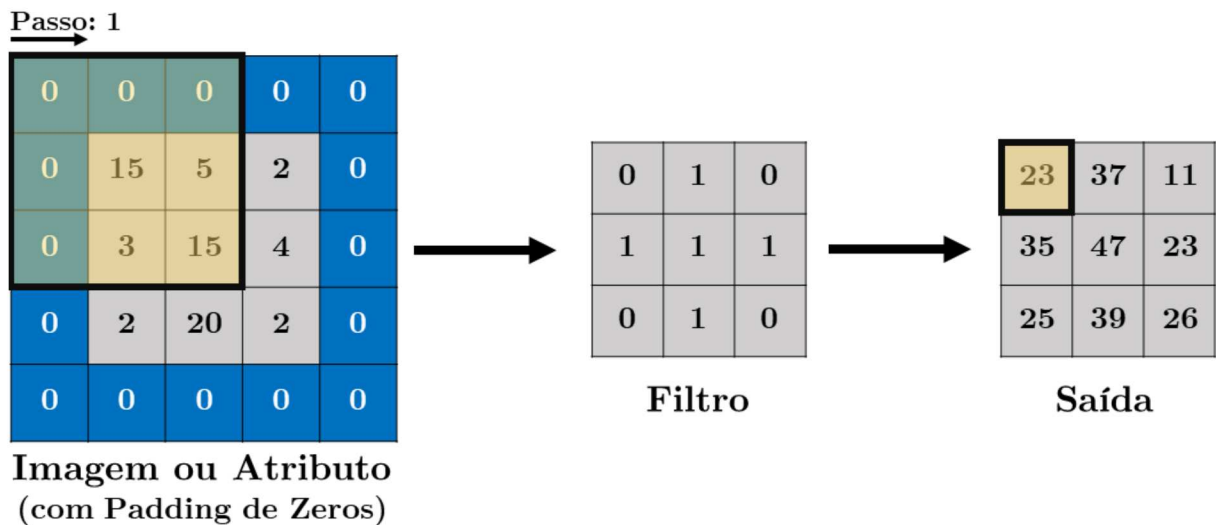
Durante o processamento, as saídas intermediárias de cada camada são denominadas mapas de atributos, que representam a informação processada até aquele estágio. Esses mapas destacam características visuais cruciais da imagem, contribuindo para uma representação mais refinada e adequada à tarefa de interesse.

As etapas de convolução desempenham um papel central no funcionamento de uma RNC. Essas etapas utilizam filtros, ou *kernels*, que percorrem os dados de entrada para extrair atributos relevantes, como bordas e texturas. Os filtros são matrizes de dimensões ajustáveis, que aplicam operações matemáticas envolvendo multiplicação e soma para gerar os mapas de atributos. O tamanho dos filtros e o número de filtros empregados são hiperparâmetros críticos, pois influenciam diretamente a capacidade da rede de identificar padrões nos dados (ALBAWI et al., 2017).

Outros hiperparâmetros importantes são o passo (*stride*) e o *padding*. O passo define o deslocamento do filtro durante a convolução, enquanto o *padding* adiciona bordas artificiais à entrada, geralmente preenchidas com zeros, para evitar a redução excessiva do tamanho do mapa de atributos (SAKIB et al., 2019).

Após a convolução, o resultado pode ser ajustado por uma camada de ativação, cuja função é introduzir não linearidades no modelo, permitindo maior expressividade na

Figura 6: Operações executadas por um filtro, utilizando *padding* de zeros e passo igual a 1.



extração de características. As funções de ativação mais utilizadas incluem a sigmoide, tangente hiperbólica (*tanh*) e a ReLU (*Rectified Linear Unit*). Segundo Nair e Hinton (2010), a ReLU possibilita um treinamento mais eficiente, preservando as informações relevantes do sinal processado.

Os mapas de atributos gerados podem passar por uma camada de *pooling*, cuja finalidade é reduzir a dimensão dos dados de forma controlada, minimizando a complexidade do modelo e atenuando o risco de sobreajuste (*overfitting*). Os métodos mais comuns de *pooling* são o *MaxPooling*, que seleciona o valor máximo de cada sub-região do mapa de atributos, e o *AveragePooling*, que calcula a média desses valores.

Na etapa final, os mapas de atributos são convertidos em vetores unidimensionais e alimentados em camadas densas, onde ocorre a combinação das características extraídas. Em aplicações de classificação, o número de saídas da camada final corresponde ao número de categorias alvo. Essas saídas passam por uma última camada de ativação, como a função *SoftMax*, que normaliza os valores para representarem probabilidades (ARIF et al., 2018).

Em resumo, as Redes Neurais Convolucionais são ferramentas robustas para o processamento de imagens, estruturadas de forma a explorar relações espaciais e identificar padrões complexos. Sua arquitetura modular e ajustável permite aplicações em diversas áreas, desde reconhecimento de objetos até diagnóstico médico assistido por imagens.

3.2.3 Modelos de Detecção de Objetos

A distinção entre modelos de detecção de objetos e modelos simples de classificação muitas vezes reside na capacidade de identificar objetos e integrar informações de lo-

calização. Enquanto modelos de classificação simples focam em atribuir uma classe ou rótulo à imagem como um todo, os modelos de detecção são projetados para identificar e localizar objetos específicos na imagem. Modelos de detecção, especialmente os baseados em aprendizado profundo, geralmente são divididos em três partes principais, conforme apresentado na Figura 7, o que facilita a compreensão da hierarquia de processamento de dados.

A espinha dorsal, comumente referida como *backbone*, é frequentemente uma Rede Neural Convolutiva (RNC) previamente treinada para servir como a base do processo de detecção de objetos. Sua função é extrair atributos da imagem de entrada de maneira hierárquica, começando com atributos de baixo nível, como bordas e texturas, em suas camadas iniciais, e progredindo para atributos de alto nível, como partes do objeto e informações semânticas e de contexto, em camadas mais profundas.

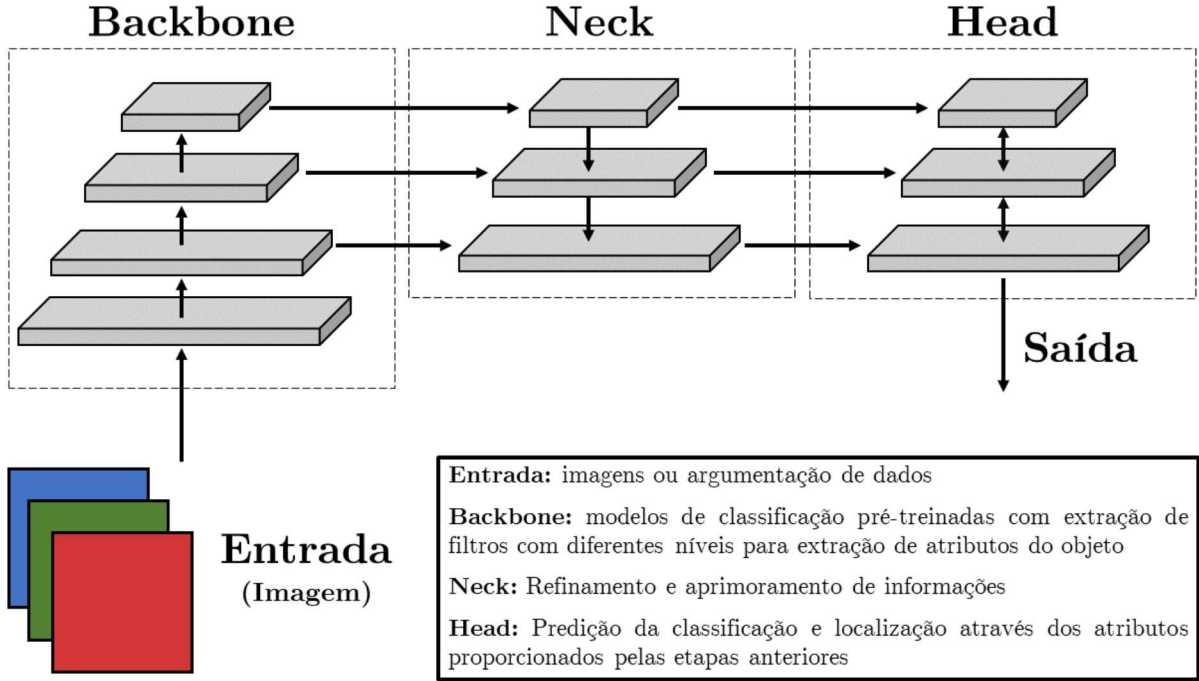
O pescoço, ou *neck*, atua como uma camada intermediária, refinando as informações provenientes da espinha dorsal e concentrando-se em aprimorar os atributos identificados. Frequentemente, essa camada incorpora camadas convolucionais adicionais ou uma rede piramidal de atributos (*Feature Pyramid Network* - FPN). Em algumas estruturas, o pescoço pode ser inexistente, tendo suas funções agregadas à espinha dorsal ou à cabeça.

A cabeça, ou *head*, representa a parte final de um detector de objetos, sendo responsável pela predição da classificação e localização com base nos atributos derivados das etapas anteriores. Essa abordagem modular, com distintas responsabilidades atribuídas a cada parte do modelo, contribui para uma compreensão mais clara e eficaz da rede de detecção de objetos.

Para avaliar a eficiência desses modelos, é crucial empregar métricas específicas que ofereçam informações sobre seu desempenho. Para isso, alguns conceitos serão introduzidos para simplificar as métricas e compreender melhor como elas influenciam a qualidade da detecção de objetos. Esses conceitos incluem:

- Verdadeiros Positivos (*TP*): quando o modelo detecta um objeto com uma caixa de identificação que contempla o objeto de interesse;
- Falsos Positivos (*FP*): quando o modelo detecta um objeto com uma caixa de identificação, mas que não contempla o objeto de interesse;
- Falsos Negativos (*FN*): quando o modelo não detecta um objeto com uma caixa de identificação, mas há um objeto de interesse na imagem;
- Precisão (*P*): é a porcentagem de *TP* dentro de todos os positivos encontrados pelo modelo, representando a capacidade do modelo de identificar apenas objetos

Figura 7: Divisão típica de modelos de detecção de objeto de aprendizado profundo.



válidos. A fórmula para precisão é dada por:

$$P = \frac{TP}{TP + FP} \quad (16)$$

- Revocação (R): é a porcentagem de TP dentro de todos os objetos verdadeiros do modelo, representando a capacidade do modelo de identificar todos os casos relevantes. A fórmula para revocação é dada por:

$$R = \frac{TP}{TP + FN} \quad (17)$$

- Interseção sobre união (IoU): é a relação entre a área de interseção entre a caixa de identificação predita pelo modelo e a caixa de identificação verdadeira para o objeto sobre a área de união entre elas. A fórmula para IoU é:

$$IoU = \frac{\text{Área de Interseção}}{\text{Área de União}} \quad (18)$$

- Limiar (l): valor mínimo atribuído ao IoU para avaliar se a detecção foi ou não correta. Se $IoU \geq l$, a detecção é considerada válida, e se $IoU < l$, é considerada inválida.

Para um modelo ser considerado bom, ambos os valores, precisão e revocação, devem

ser o mais próximo possível de 1, indicando que o modelo tem o mínimo de falsos positivos (tornando a precisão mais alta, identificando apenas objetos válidos) e falsos negativos (tornando a revocação mais alta, identificando todos os objetos relevantes) (ZOU et al., 2023).

Para modelos de detecção, a métrica comumente utilizada é a precisão média (PM), que avalia o desempenho da localização e da classificação de um objeto para um modelo. Esse valor pode ser avaliado pela análise gráfica da relação precisão versus revocação, sendo seu valor dado pela área abaixo da curva. A integral para calcular a precisão média pode ser aproximada usando métodos como a interpolação de 11 passos (SCHÜTZE et al., 2008). A fórmula para o cálculo da precisão média é:

$$PM = \int_0^1 P(R) dR \quad (19)$$

Contudo, a precisão média se dá para apenas um objeto. Sendo assim, para modelos que possuem múltiplas classificações, utiliza-se a média das precisões médias (mAP).

$$mAP = \frac{1}{N} \sum_{i=1}^N PM_i \quad (20)$$

sendo i o índice da classificação e N o número de classes presentes no modelo.

Além disso, modelos de detecção podem ser divididos em dois grupos principais: os de estágio duplo e os de estágio único. Ambos os tipos de modelos utilizam a estrutura espinha dorsal-pescoço-cabeça para organizar suas arquiteturas.

Os modelos de estágio duplo propõem a integração de duas Redes Neurais Convolucionais (RNCs) para a detecção de objetos. Na primeira etapa, a espinha dorsal (ou *backbone*) é responsável por extrair atributos da imagem de entrada e identificar possíveis regiões de interesse (RoIs), que são possíveis caixas delimitadoras onde os objetos podem estar localizados. Essas RoIs são então refinadas pelo pescoço (*neck*), que frequentemente incorpora camadas convolucionais adicionais ou uma Rede Piramidal de Atributos (*Feature Pyramid Network* - FPN), concentrando-se em aprimorar os atributos identificados para cada região de interesse. Na segunda etapa, a cabeça (*head*) da rede é encarregada de classificar o objeto contido dentro das caixas delimitadoras refinadas e ajustar a localização das caixas delimitadoras com maior precisão.

Um exemplo clássico de modelo de estágio duplo é a R-CNN (Regional Convolutional Neural Network) e suas variantes, como a Fast R-CNN e a Faster R-CNN. Na Faster R-CNN, por exemplo, a espinha dorsal é geralmente uma rede profunda como a ResNet ou VGG, que extrai mapas de atributos da imagem. Uma Rede de Proposta de Regiões (*Region Proposal Network* - RPN) é então utilizada para gerar propostas de caixas delimitadoras, que são refinadas e classificadas em etapas subsequentes (REN et

al., 2015).

Por outro lado, os modelos de estágio único realizam o processo de detecção usando apenas uma RNC, que desempenha simultaneamente a localização e a classificação dos objetos. Nesse caso, a espinha dorsal, o pescoço e a cabeça são combinados em uma única rede, simplificando a arquitetura e acelerando o processo de detecção. Esses modelos predizem diretamente as classes e as caixas delimitadoras para várias escalas e localizações na imagem de entrada em uma única etapa de processamento.

Um exemplo popular de modelo de estágio único é o YOLO (You Only Look Once) e suas variantes. No YOLO, a imagem é dividida em uma grade, e para cada célula da grade, a rede prevê diretamente as classes dos objetos presentes e os offsets das caixas delimitadoras relativas àquela célula. Outro exemplo é o SSD (Single Shot MultiBox Detector), que utiliza múltiplas escalas e aspectos de caixas delimitadoras em diferentes níveis de uma rede convolucional para realizar a detecção em uma única etapa.

Ambos os tipos de modelos, de estágio duplo e de estágio único, utilizam a estrutura espinha dorsal-pescoço-cabeça para organizar suas arquiteturas. Esta abordagem modular permite uma clara separação de responsabilidades em cada parte do modelo, facilitando a compreensão e o treinamento eficaz para a detecção de objetos.

A distinção inicial entre esses métodos está ancorada nas arquiteturas, as quais, por sua vez, priorizam enfoques diferenciados entre precisão e velocidade de predição do modelo. Em geral, os modelos de estágio duplo tendem a oferecer maior precisão, visto que cada uma das redes se especializa em uma ação específica, contribuindo para um treinamento mais refinado dos dados de entrada e saída. Essa abordagem revela-se eficiente na identificação tanto da presença quanto da ausência de objetos, além de proporcionar uma classificação precisa dos objetos identificados (LU et al., 2020).

Em contrapartida, o modelo de estágio único é treinado em apenas uma rede, resultando na simultânea localização e classificação. Isso torna seu treinamento mais complexo, uma vez que há a necessidade de ajustar pesos e bias para duas saídas correlacionadas em uma única entrada, dificultando a obtenção de um resultado com alta precisão. No entanto, devido à sua abordagem singular, ao processar novas imagens com os parâmetros já obtidos após o treinamento, ele demonstra ser mais rápido.

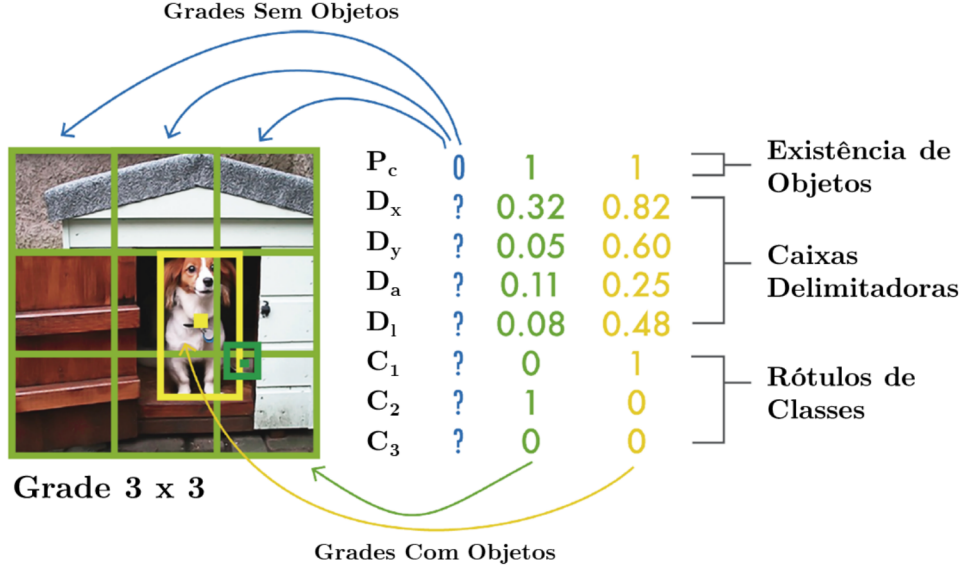
3.2.4 Algoritmo de Detecção de Objetos YOLO

Ao longo dos anos, diversos modelos de detecção foram propostos e aprimorados. Entre esses modelos, destaca-se o YOLO, uma abreviação para "You Only Look Once" (em tradução livre, "Você Só Olha uma Vez- YOLO). O algoritmo foi introduzido por Joseph Redmon em 2015, propondo uma abordagem de estágio único para detecção de objetos em tempo real.

Para melhorar o desempenho do algoritmo, Redmon propôs unificar vários passos

da detecção, detectando todas as caixas delimitadoras simultaneamente. A imagem de entrada é dividida em grades de tamanho variável $M \times M$, onde cada grade prediz D caixas de detecção, juntamente com a confiança para cada uma das C classes.

Figura 8: Exemplo de análise por grades utilizando o YOLO, adaptado de Terven et al. (2023).



Cada grade detecta uma caixa delimitadora, associada a 5 variáveis: o valor de confiança da caixa P_c , representando a probabilidade de realmente haver um objeto; as coordenadas do centro da caixa (D_x e D_y); e suas dimensões (D_a e D_l). Além disso, a grade também prediz um vetor contendo o valor de confiança para cada uma das C classes.

O algoritmo retorna um tensor de tamanho $M \times M \times (B \times 5 + C)$. Esse resultado é frequentemente seguido pelo supressor de não-máximos (SNM), um algoritmo que age como um filtro, visando reduzir as caixas encontradas pelo algoritmo que se sobrepõem e, muitas vezes, são geradas ao redor do mesmo objeto. O SNM mantém apenas as caixas de maior acurácia.

Na primeira versão, o algoritmo propôs uma arquitetura de 24 camadas convolucionais de dimensão 1×1 , com algumas camadas de *pooling* utilizando *MaxPooling* como operação, seguidas por duas camadas totalmente conectadas. Todas foram seguidas por uma camada de ativação ReLU, exceto a última, que utilizava uma função de ativação linear (REDMON et al., 2016).

No entanto, embora sua velocidade fosse superior a outros métodos de detecção de ponta na época, como o *Fast R-CNN* (GIRSHICK, 2015), o modelo apresentava limitações. Entre elas, a incapacidade de detectar mais de dois objetos da mesma classe em uma mesma grade, treinamento através de atributos mais amplos dos objetos devido à redução das amostragens nas camadas, diminuindo a capacidade de detectar objetos

com mais detalhes, e dificuldade em prever objetos com proporções muito diferentes dos dados de treinamento (TERVEN et al., 2023).

Desde então, a arquitetura foi recebendo atualizações com o intuito de mitigar essas problemáticas e melhorar o desempenho do algoritmo, tanto na precisão média quanto na velocidade de processamento.

- **YOLO V2:** Apresentado em Redmon e Farhadi (2017), teve como foco a manutenção da velocidade do modelo anterior, mas permitindo agora detectar 9000 categorias. As maiores alterações foram tornar a rede totalmente convolucional, substituindo a camada densa totalmente conectada; utilização do método de ancoragem para prever as caixas delimitadoras; e alteração da estrutura de classificação para DarkNet-19. Essas alterações permitiram, para os mesmos dados de treinamento, promover um MPM de 78,6% em comparação com 63,4%, apresentando uma melhora significativa no desempenho do modelo. Sua limitação maior se encontrava na dificuldade em detectar objetos com proporções.
- **YOLO V3:** Redmon e Farhadi (2018) apresentou o desenvolvimento de uma nova arquitetura denominada Darknet-53, que apresenta 53 camadas convolucionais, substituindo todas as operações de *MaxPooling* com convoluções utilizando passos diferentes e a adição de redes neurais residuais (ResNet). Esta permite que redes profundas não percam informações, criando "atalhos" com esses importantes atributos permitindo que eles não sejam perdidos ao longo do treinamento. Outra proposição foi a utilização de predição em múltiplas escalas, que permite realizar o treinamento utilizando múltiplos tamanhos de grades, mitigando a limitação para objetos pequenos apresentada pelo seu predecessor.

Utilizando um banco de dados mais robusto, denominado *Microsoft COCO* (LIN et al., 2014), foi avaliado o desempenho comparativo entre a segunda e terceira versões, obtendo respectivamente as MPM de 21,6% e 33,0%. Além disso, a terceira versão apresentou uma velocidade de processamento duas vezes mais rápida, demonstrando a melhora no desempenho na versão.

- **YOLO V4:** Bochkovskiy et al. (2020) propôs algumas alterações estruturais com o objetivo de aprimorar o desempenho da rede proposta por Redmon e Farhadi (2018). Alterou a tradicional função ReLU pela função de ativação Mish (MISRA, 2019), que se apresentou com uma função de auto-regularização, evitando *overfitting*, e pela sua suavidade e monotonicidade, permitindo treinamentos mais estáveis e constantes. Houve também um balanceamento entre conjuntos de recursos especiais (*Bag of Specials*) e o conjunto de brindes (*Bag of Freebies*). O primeiro refere-se a métodos para melhorar o treinamento do modelo, sem aumentar o tempo de processamento, sendo feito, nesse caso, através da avaliação da parte estrutural da

arquitetura. O segundo se refere a métodos que apresentam melhorias na precisão do modelo, muitas vezes aumentando o tempo de inferência do treinamento. Para esse modelo, tem-se a utilização da composição de um mosaico combinando mais de uma imagem, tornando essa uma das entradas do modelo e reduzindo a quantidade de bateladas necessárias para treinar o modelo, e a introdução do treinamento auto-adversarial (SAT), distorcendo a imagem original tornando o objeto menos visível, mas mantendo a caixa de detecção de entrada, permitindo que o modelo se torne mais robusto a variações e encontre mais rápido padrões.

Utilizando o mesmo banco de dados de dados, *Microsoft COCO*, o algoritmo apresentou um MPM de 43,5%, superior a seu predecessor, porém com um tempo de execução considerado muito semelhante e um maior esforço computacional.

- **YOLO V5:** A quinta versão se apresentou pela primeira vez vinculada a uma empresa, sendo ela a Ultralytics, e se tornando uma biblioteca aberta tornando-se amplamente utilizável. O novo modelo, apresentado em Jocher et al. (2020), ao contrário das predecessores deixou de utilizar o framework Darknet para as camadas convolucionais, optando pelo Pytorch para possibilitar a implementação de outro algoritmo da mesma empresa, o AutoAnchor, que permite encontrar âncoras nas caixas delimitadoras de entrada de maneira automática, tornando mais simplificada a manipulação dos dados, sem ser necessário avaliar tais âncoras de maneira manual. O algoritmo sofreu ajustes finos em sua arquitetura, que apresentaram melhor desempenho, porém manteve-se muito semelhante a sua versão anterior. Contudo introduziram a capacidade de escolher 5 escalonamentos distintos que alteram que alteram as dimensões dos módulos convolucionais, alterando conforme a necessidade do usuário (capacidade da máquina e dos dados utilizados). Em ordem crescente de desempenho e necessidade de hardware, são elas *Nano* (*n*), *Small* (*s*), *Medium* (*m*), *Large* (*l*), e *Extra-large* (*x*). Sendo capaz de alcançar 55,8% de MPM com os dados *Microsoft COCO* utilizando a versão mais robusta *Extra-large*.
- **YOLO V6:** Baseado na arquitetura YOLO v4, Li et al. (2022) propôs a utilização de uma estrutura de *EfficientNets* (TAN; LE, 2019), que utiliza escalonamento para reduzir a rede mas sem prejudicar sua eficiência. Além disso, propõem o método de ancoragem livre, portanto prevê diretamente a caixa delimitadora sem a necessidade de avaliar âncoras pré-determinadas. Assim como o YOLO V5, apresenta escalas distintas, possuindo oito no total, apresentando um MPM 57,2%.
- **YOLO V7:** Proposto pelo mesmo autor da YOLOV4, o modelo apresentado em Wang et al. (2023) utiliza re-parametrização planejada durante as convoluções, de maneira a ajustar os parâmetros de maneira que o caminho de propagação dos

resultados seja otimizado, aperfeiçoando e otimizando o mesmo. Utilizou-se de escalonamento para modelos baseados em concatenação, implicando que as dimensões das camadas se manterão proporcionais entre si, podendo escaloná-las conforme a necessidade sem alterar a estrutura ótima proposta. Por fim, propôs a utilização de blocos de rede de agregação eficiente de camadas estendidas (E-ELAN), que utiliza técnicas como expansão, embaralhamento e fusão de cardinalidade para aprimorar continuamente a capacidade de aprendizado da rede sem comprometer o caminho de gradiente original, o que permite um aprendizado mais eficiente (WANG et al., 2022). Com essas alterações, o seu escalonamento mais robusto foi capaz de alcançar um MPM de 55,9%.

- **YOLO V8:** Disponibilizada em janeiro de 2023 em Jocher et al. (2023), a oitava versão é um aprimoramento da YOLOV5, sendo da mesma empresa, Ultralytics. As principais mudanças em comparação à quinta versão incluem a adoção do método de ancoragem livre e o uso da nova camada convolucional denominada C2f, um módulo contendo duas convoluções entre estágios que combinam atributos de alto nível com informações contextuais. A versão *Extra-large* alcançou 57% de MPM com os dados *Microsoft COCO*, com uma velocidade de processamento de imagem de até 280FPS.

Dessa forma, as evoluções do modelo permitiram que o YOLO sempre alcançasse o estado da arte a cada nova versão. Ele é equiparado, em níveis de precisão, a outros modelos consolidados na detecção de imagens, como Faster R-CNN (GIRSHICK, 2015) e SSD - Detector de Multi Caixas por Disparo Único (LIU et al., 2016). No entanto, o YOLO se destaca por sua eficiência em termos de velocidade de processamento, tornando-se mais capaz de realizar detecções em tempo real.

3.2.5 Método de Fluxo Óptico

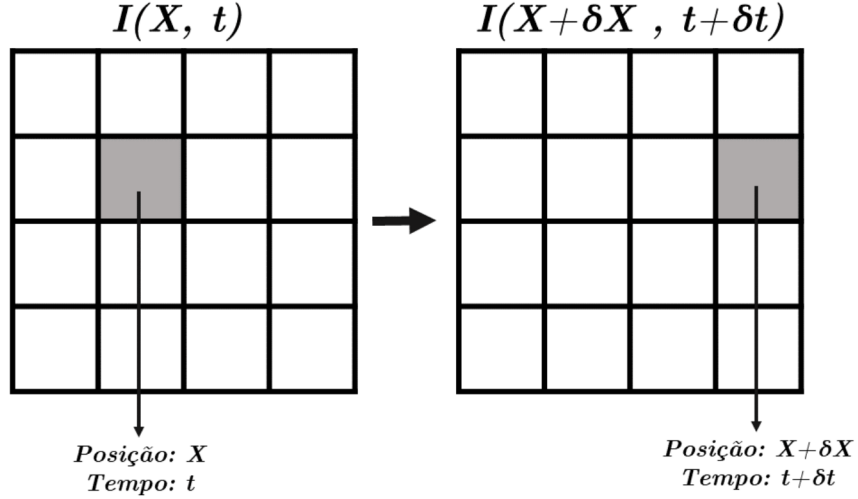
O método de fluxo óptico é uma técnica amplamente utilizada no rastreamento de objetos, baseada na suposição de que a intensidade luminosa dos pixels de um objeto em uma imagem permanece aproximadamente constante ao longo do tempo. O método busca detectar e interpretar variações na intensidade dos pixels entre quadros consecutivos, identificando conjuntos de pixels com intensidades semelhantes (BEAUCHEMIN; BARRON, 1955).

Matematicamente, a intensidade do pixel em uma imagem ($I(X, t)$) pode ser aproximada como:

$$I(X, t) \approx I(X + \delta X, t + \delta t) \quad (21)$$

Isso implica que um ou mais pixels locais que compõem um objeto em movimento apresentarão intensidades semelhantes mesmo após um pequeno deslocamento (δX) em um curto período de tempo (δt), como apresentado esquematicamente na Figura 9.

Figura 9: Esquema de pontos de mesma intensidade após um deslocamento temporal δt com um deslocamento δX .



Utilizando a série de Taylor truncada até o segundo termo, a Equação 22 pode ser reescrita como:

$$I(X + \delta X, t + \delta t) \approx I(X, t) + \nabla I \cdot \delta X + \delta t I_t + \dots \quad (22)$$

onde $\nabla I = (I_x, I_y)$ e I_t são as derivadas parciais de primeira ordem de $I(X, t)$. Truncando a série no segundo termo, substituindo-a na Equação 21, subtraindo $I(X, t)$ de ambos os lados e dividindo por δt , obtemos a Equação 22:

$$\nabla I \cdot V + I_t = 0 \quad (23)$$

Na Equação 23, $V = (u, v)$ representa a velocidade dos pixels entre um quadro e outro. O objetivo é encontrar um vetor de velocidade mínima que satisfaça a Equação 23, o que só é possível se houver uma concentração adequada de intensidades semelhantes. Isso implica que objetos muito pequenos (pequenas intensidades de pixels) ou com texturas altamente heterogêneas (pixels de intensidade distintas) podem não ser detectados com precisão pelo algoritmo de fluxo óptico.

Para que o método de fluxo óptico seja aplicável de maneira eficiente, e portanto seja capaz de caracterizar e rastrear com maior precisão o deslocamento dos conjuntos luminosos entre os quadros, é necessário que algumas condições sejam atendidas:

- Iluminação uniforme: a intensidade luminosa do ambiente na captura das imagens não deve variar significativamente;

- Superfície com reflectância lambertiana: a intensidade da imagem não deve depender do ângulo de visão do observador;
- Movimento de translação pura e paralela ao plano da imagem: o movimento do objeto ocorre apenas paralelamente ao plano da imagem.

Portanto, a aplicação do método de fluxo óptico envolve a análise de variações na intensidade dos pixels em uma imagem para detectar objetos em movimento e rastreá-los ao longo dos quadros. Isso permite a estimativa dos vetores de movimento entre os pixels, considerando as intensidades de brilho (I) nas direções horizontal e vertical (x, y) ao longo do tempo (t).

Esse tipo de abordagem pode ainda ser dividida em duas subcategorias: os métodos de fluxo óptico local e global. De maneira geral, uma maneira comum de entender a diferença entre os métodos locais e globais é pensar na escala da análise.

Os métodos globais, representados por técnicas como a proposta de Horn-Schunck (HORN; SCHUNCK, 1981), consideram informações em toda a cena para estimar o movimento. Eles analisam a imagem como um todo e buscam entender o contexto geral do movimento, levando em conta as interações entre todos os pixels. Essa abordagem é vantajosa em cenários onde há variações complexas de movimento em toda a cena, como em vídeos de tráfego urbano ou em ambientes naturais. Os métodos globais são mais robustos a ruídos locais e oclusões parciais, pois consideram o contexto completo da imagem ao calcular o fluxo óptico. No entanto, eles podem exigir mais recursos computacionais do que os métodos locais, o que pode ser uma consideração importante em aplicações com restrições de tempo ou recursos.

Por outro lado, os métodos locais, representados por técnicas como o método de Lucas-Kanade (LUCAS; KANADE, 1981), focam na análise de regiões específicas da imagem para estimar o movimento localmente. Esse método assume que o movimento dentro dessas regiões pode ser aproximado por um vetor de deslocamento constante. Isso significa que ele calcula o fluxo óptico independentemente em cada região, sem considerar o contexto global da cena. Essa abordagem é vantajosa em cenários onde o movimento é relativamente homogêneo e previsível dentro de pequenas áreas da imagem, como em rastreamento de objetos em um ambiente controlado. No entanto, os métodos locais podem ser menos eficazes em situações com movimentos variáveis, presença de ruídos ou oclusões parciais, onde a suposição de movimento constante pode ser violada. Essa técnica é particularmente útil para estudar o comportamento em condições controladas e com movimentação previsível.

O método de Lucas-Kanade (LK) é uma escolha vantajosa para medir a velocidade em bolhas alongadas. Sua capacidade de rastrear grandes deslocamentos de forma robusta e detectar variações sutis na velocidade ao longo da bolha é essencial nesse contexto, como

mostram os vetores de deslocamento na Figura 10, onde os pontos rosa representam os pixels avaliados e as linhas em verde indicam o deslocamento desses pontos ao longo dos últimos 5 quadros. Além disso, sua eficiência computacional permite sua aplicação em tempo real ou em sistemas com recursos limitados. Este método também lida bem com a diversidade de texturas na superfície da bolha, comum nesse tipo de fenômeno. Outra vantagem é sua implementação direta usando a biblioteca OpenCV, que oferece uma interface acessível e ferramentas robustas para análise de resultados.

Figura 10: Vetores de deslocamento por pixel formados através do método de fluxo óptico Lucas-Kanade, na ponta da bolha de Taylor.



Isso torna o algoritmo mais eficiente computacionalmente e adequado para aplicações de rastreamento de movimento em tempo real. Para essa seleção, são frequentemente utilizados métodos de detecção de cantos, como o método proposto por Shi et al. (1994), desta maneira pode-se utilizar dos contornos como os pixels de interesse.

O método é uma extensão do algoritmo de detecção de cantos de Harris (HARRIS

et al., 1988). Enquanto o algoritmo de Harris considera a matriz de covariância dos gradientes locais para cada ponto na imagem, o método de Shi-Tomasi utiliza o menor autovalor da matriz de covariância para determinar a qualidade do canto.

Para cada ponto candidato a canto na imagem, o método de Shi-Tomasi calcula a matriz de covariância dos gradientes locais, representada por M , em uma região ao redor do ponto. Esta matriz é dada por:

$$M = \sum_{i,j} w(x_i - x, y_j - y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (24)$$

onde I_x e I_y são as derivadas parciais da intensidade da imagem em relação às coordenadas x e y , respectivamente, e w é uma função de ponderação que atribui maior importância aos gradientes próximos ao ponto de interesse.

Em seguida, o método de Shi-Tomasi calcula os autovalores k_1 e k_2 da matriz M . O menor desses autovalores, denotado por R , é utilizado como medida da qualidade do canto. Quanto maior o valor de R , mais provável é que o ponto seja um canto.

Portanto, o método de Shi-Tomasi seleciona como cantos os pontos onde R é localmente máximo em relação aos seus vizinhos. Isso permite identificar cantos com uma medida robusta de sua qualidade, levando em consideração a distribuição dos gradientes locais ao redor do ponto de interesse.

Com esses pontos o método pretende resolver a equação o problema de um sistema de equações 23. Para um sistema bidimensional onde $X = (x, y)$ a equação pode ser apresentada como:

$$I_x \cdot u + I_y \cdot v = -I_t \quad (25)$$

onde I_x e I_y são as derivadas parciais da intensidade dos pixels em relação às direções x e y , respectivamente; u e v são as componentes do vetor de movimento a serem calculadas; e I_t é a variação temporal da intensidade.

Como a Equação 25 apresenta apenas a relação de um pixel, a composição de todo o conjunto de pixels de mesma intensidade da dada por toda a imagem, pode ser expressa por:

$$\begin{aligned} I_x(p_1)u + I_y(p_1)v &= -I_t(p_1) \\ I_x(p_2)u + I_y(p_2)v &= -I_t(p_2) \\ &\vdots \\ I_x(p_m)u + I_y(p_m)v &= -I_t(p_m) \end{aligned} \quad (26)$$

Que pode ser isolado formando:

$$\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_n) & I_y(p_n) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -I_t(p_1) \\ -I_t(p_2) \\ \vdots \\ -I_t(p_n) \end{bmatrix} \quad (27)$$

onde p_i represente o pixel unitário dos n pixels de mesma intensidade entre quadros.

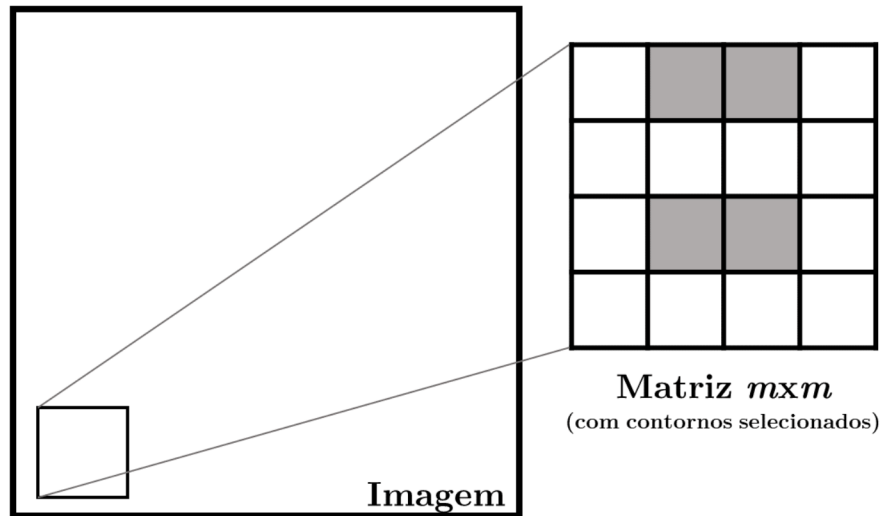
Para resolver o problema de um sistema de equações sobredeterminado, é utilizado o método dos mínimos quadrados ordinários para encontrar as melhores estimativas de movimento. Isso pode ser representado pela seguinte equação matricial:

$$\begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix} \quad (28)$$

Com base nas derivadas parciais da intensidade dos pixels e na variação temporal, o algoritmo calcula as componentes do vetor de movimento u e v para cada ponto de interesse. Essas estimativas representam o deslocamento local dos pontos na imagem.

Como apresentado anteriormente, uma das características do método LK é a adição da suposição de "consistência espacial", que implica que os pixels vizinhos em uma determinada região, dentro da imagem, têm variações semelhantes. Essa suposição ajuda a melhorar a precisão da estimativa de movimento ao considerar a relação entre os pixels próximos. Com base nessa suposição, uma pequena janela de tamanho $m \times m$ pode ser desenhada ao redor dos pontos de interesse e determina-se que todos os pontos de pixel dentro da pequena janela têm o mesmo movimento, como apresentado na Figura 11 (SU et al., 2022).

Figura 11: Apresentação figurativa da consistência espacial proposta por Lucas-Kanade.



O método de fluxo óptico LK é fundamentado na premissa de que o movimento do

objeto entre quadros consecutivos é pequeno. Essa premissa permite que a expansão de Taylor seja truncada até o primeiro termo durante o cálculo das equações de restrição, simplificando assim a estimativa do vetor de movimento. Movimentos pequenos garantem que a mudança local e global na distribuição de pixels na imagem varie pouco, resultando em estimativas mais precisas com o LK.

Por outro lado, quando o movimento entre quadros é grande, a intensidade dos pixels pode variar consideravelmente, o que compromete a suposição de que a intensidade luminosa dos pixels é constante. Isso pode levar a grandes erros na estimativa do vetor de movimento utilizando apenas a expansão de Taylor truncada. Para resolver esse problema, o método de fluxo óptico LK pode ser aprimorado utilizando o método de pirâmide.

O método de pirâmide consiste em dividir a imagem em várias escalas (ou níveis da pirâmide). No nível mais baixo da pirâmide, a imagem é reduzida para uma resolução menor, o que reduz o movimento aparente entre os quadros consecutivos. O vetor de movimento é estimado incrementalmente em cada nível da pirâmide, começando pelo nível mais baixo e subindo até o nível mais alto. Esse processo permite que o método lide eficazmente com grandes deslocamentos, melhorando a precisão da estimativa final do vetor de movimento.

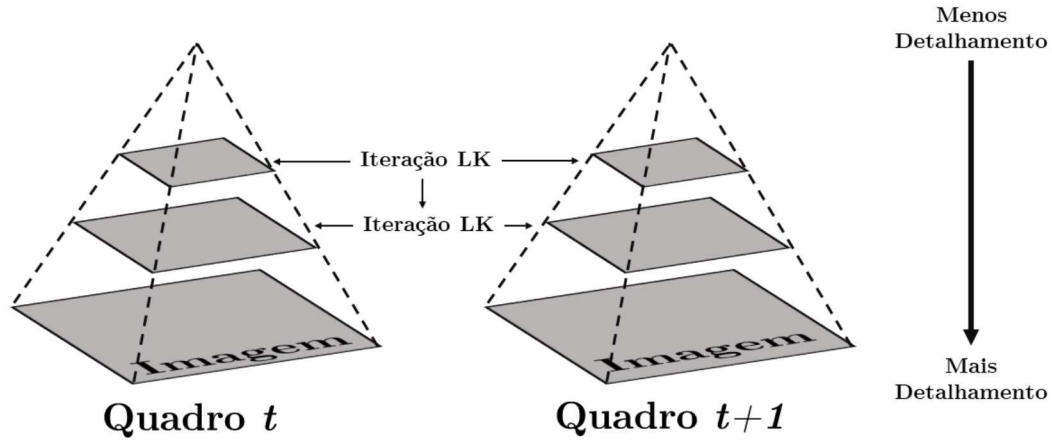
Portanto, enquanto o método de LK é eficaz para pequenos movimentos devido à estabilidade na distribuição de pixels entre quadros consecutivos, o método de pirâmide é essencial para lidar com grandes movimentos, refinando a estimativa do vetor de movimento ao longo das escalas da pirâmide.

Como demonstrado na Figura 12, ao amostrar a imagem, o método de fluxo óptico LK em pirâmide pode reduzir os deslocamentos maiores na imagem de pirâmide de nível superior para obter um vetor de fluxo óptico mais preciso nessa escala e, em seguida, ampliar o vetor de nível superior como um guia inicial para a próxima camada ao resolver camada por camada de cima para baixo. Neste momento, a posição do objeto guiada pelo vetor de fluxo óptico do nível superior ampliado.

Isso ajuda a evitar o bloqueio de movimento (*motion blocking*), que é um fenômeno onde o algoritmo não consegue rastrear objetos que se movem mais do que um determinado limite entre os quadros consecutivos.

Em resumo, o método de fluxo óptico Lucas-Kanade, integrado ao método Shi-Tomasi e o processo de piramidação, é uma técnica viável para rastreamento de movimento em vídeos. Ele opera sobre a suposição de que o movimento na cena entre dois quadros consecutivos é pequeno e, portanto, pode ser aproximado por um vetor de movimento constante em uma pequena vizinhança de pixels. Utilizando o método dos mínimos quadrados, o algoritmo calcula o vetor de movimento que melhor se ajusta às restrições de luminosidade entre os quadros. O método pode ser aplicado a uma variedade de aplicações, desde rastreamento de objetos em vídeos de vigilância até análise de

Figura 12: Demonstração do processo de piramidação do método de fluxo óptico Lucas-Kanade.



movimento em vídeos médicos.

3.3 Métodos Estatísticos

Nesta seção, serão apresentados os métodos estatísticos que serão utilizados ao longo desta dissertação.

3.3.1 Média e Desvio Padrão

A média e o desvio padrão são medidas estatísticas fundamentais usadas para descrever e entender conjuntos de dados na estatística e na análise de dados.

A média aritmética, é uma medida de centralidade que representa o valor típico ou central em um conjunto de dados. Para calcular esta média, soma-se todos os valores nos dados e divide-se pelo número total de observações. A média é calculada pela fórmula:

$$\mu = \frac{\sum_{i=1}^n x_i}{n_{\text{tot}}} \quad (29)$$

onde μ é a média, x_i são os valores individuais no conjunto de dados e n_{tot} é o número total de observações.

A média é sensível a valores extremos, o que significa que um único valor muito grande ou muito pequeno pode distorcer significativamente o resultado. No entanto, é uma medida útil para entender o valor típico de um conjunto de dados.

O desvio padrão é uma medida de dispersão que indica o quanto os valores de um conjunto de dados tendem a se afastar da média. Ele fornece uma medida da variabilidade ou dispersão dos dados. Quanto maior o desvio padrão, mais dispersos estão os dados em relação à média.

O desvio padrão é calculado tomando a raiz quadrada da variância. A variância é

a média dos quadrados das diferenças entre cada valor e a média. Matematicamente, o desvio padrão é calculado pela fórmula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n_{\text{tot}}}} \quad (30)$$

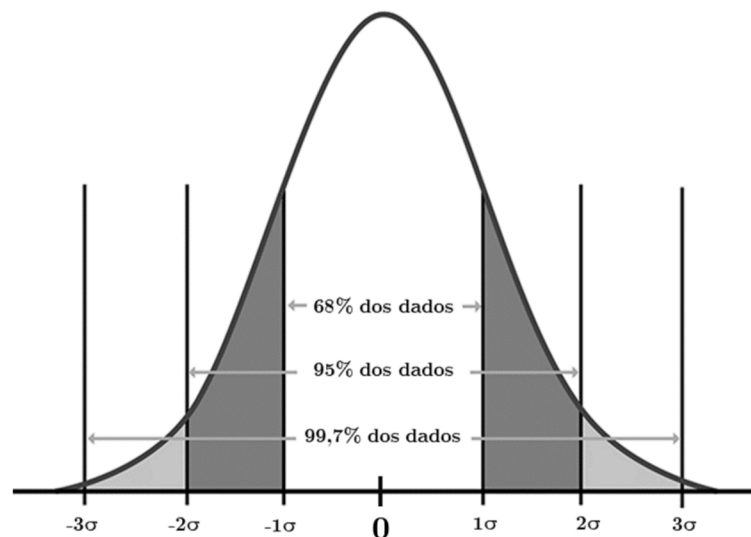
onde σ é o desvio padrão e x_i são os valores individuais no conjunto de dados.

Conjuntos de dados com desvio padrão maior têm mais variação, enquanto aqueles com desvio padrão menor têm menos variação.

Compreender a relação entre o desvio padrão e a distribuição normal é fundamental para a estatística e a análise de dados. Na distribuição normal, também conhecida como distribuição de Gauss ou distribuição de sino, os dados são simetricamente distribuídos em torno da média, formando uma curva em forma de sino.

O desvio padrão desempenha um papel crucial na distribuição normal, pois ele determina a dispersão dos dados em torno da média e, conseqüentemente, a forma e a largura da curva da distribuição normal. Na distribuição normal, cerca de 68% dos dados estão dentro de um desvio padrão da média, aproximadamente 95% dos dados estão dentro de dois desvios padrão da média, e aproximadamente 99,7% dos dados estão dentro de três desvios padrão da média (LIVINGSTON, 2004). A distribuição dos dados pode ser visualizada na Figura 13.

Figura 13: Curva normal com apresentação dos desvios padrões, considerando 0 a média.

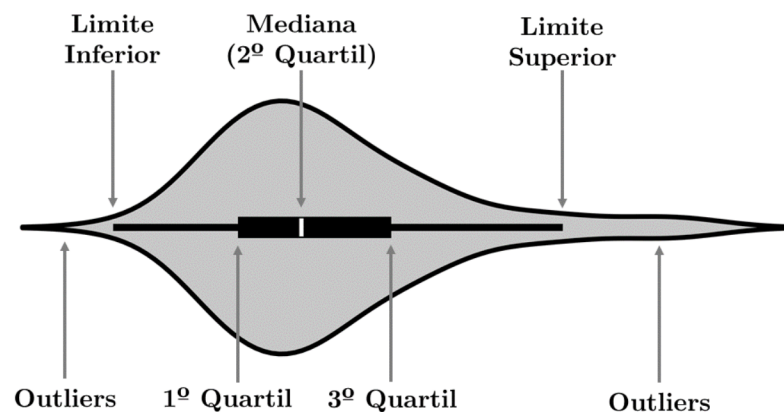


Em resumo, a média fornece uma medida central dos dados, enquanto o desvio padrão oferece uma medida da dispersão desses dados em torno da média. Ambas as medidas são cruciais para entender a distribuição e a variabilidade dos dados em uma análise estatística.

3.3.2 Distribuição em Violino

Para avaliar os dados, foi realizada uma representação visual da dispersão desses dados no formato de violino. O método do violino é uma forma de visualização de dados apresentada inicialmente por Hintze e Nelson (1998), que combina elementos do diagrama de caixa (*box plot*) e do gráfico de densidade de probabilidade para mostrar a distribuição dos dados em relação a diferentes categorias ou grupos. Cada "violino" no gráfico mostra a distribuição dos dados para uma categoria específica, com a altura indicando a densidade dos dados. Marcas individuais podem representar valores de dados individuais. O contorno do violino é uma estimativa da densidade de probabilidade dos dados, calculada usando a Estimação de Densidade do Núcleo (*Kernel Density Estimation - KDE*), que suaviza a distribuição dos dados. Em resumo, o gráfico de violino oferece uma representação visual rica e informativa da distribuição dos dados e suas variações em diferentes grupos ou categorias.

Figura 14: Esquema de apresentação de dados em violino



Na Figura 14, temos a apresentação gráfica de um violino, onde pode-se visualizar o segundo quartil (2Q), representado pela mediana dos dados, os primeiros (1Q) e terceiros (3Q) quartis que representam, respectivamente, 25% dos dados abaixo e acima da mediana. Assim, a região entre os quartis apresenta 50% de todo o conjunto de dados (BLAND, 2015).

A mediana, por sua vez, é uma medida de centralidade que representa o valor do meio de um conjunto de dados quando organizado em ordem crescente ou decrescente, dividindo o conjunto em duas partes iguais. Diferentemente da média, a mediana não é afetada por valores extremos ou discrepantes, tornando-a uma medida robusta de tendência central em situações onde os dados podem ser influenciados por valores atípicos. Além disso, há a existência dos limites superior e inferior, que separam os dados dos *outliers*, que são pontos considerados muito afastados do restante da distribuição.

3.4 Conclusão

A fundamentação teórica apresentada ao longo desta seção estabeleceu os alicerces necessários para a compreensão e análise do problema de estudo, fundamentando as escolhas metodológicas e teóricas adotadas. Foram abordados conceitos essenciais sobre escoamentos bifásicos, destacando a importância da medição precisa das interações entre as fases líquida e gasosa em setores industriais cruciais, como o petroquímico e o de energia nuclear. A complexidade desses escoamentos exige metodologias de medição que superem as limitações das abordagens convencionais, especialmente no que se refere à precisão e capacidade de lidar com dados dinâmicos.

A introdução de técnicas de visão computacional, especialmente com o uso das redes neurais convolucionais (YOLOv8) e do método de fluxo óptico Lucas-Kanade, foi central para a construção do arcabouço teórico. Essas abordagens permitem a detecção não intrusiva das bolhas, oferecendo uma solução inovadora que potencializa a análise de escoamentos líquido-gás em condições complexas, o que representa um avanço em relação às técnicas tradicionais.

Além disso, a análise estatística, que envolveu métodos como média, desvio padrão e gráficos de violino, contribuiu para a organização e interpretação dos dados, fornecendo uma base robusta para validar os resultados experimentais. O emprego dessas ferramentas permitiu uma compreensão mais profunda das variações e tendências nas medições, oferecendo uma visão detalhada da dinâmica dos escoamentos.

Em conjunto, as abordagens computacionais e estatísticas discutidas, aliadas ao entendimento dos escoamentos bifásicos, reforçam a hipótese de que a análise da velocidade das bolhas de Taylor, por meio de técnicas computacionais, pode oferecer uma alternativa eficaz às metodologias convencionais, proporcionando maior precisão e flexibilidade nas medições.

Portanto, os conceitos abordados nesta seção fundamentam de maneira sólida a metodologia adotada neste estudo, preparando o terreno para as etapas subsequentes, nas quais serão apresentados os métodos experimentais e os resultados obtidos. A análise teórica não apenas contextualiza o problema, mas também oferece os direcionamentos necessários para a exploração da metodologia e a validação dos objetivos do trabalho.

4 Revisão Bibliográfica

Esta seção apresenta uma revisão da literatura sobre estudos anteriores que investigaram o comportamento das bolhas de Taylor em escoamentos bifásicos verticais por meio de métodos de visão computacional. Na seção 4.1, são discutidos os estudos que empregaram modelos baseados em redes neurais convolucionais para a classificação e detecção de escoamentos bifásicos, com ênfase nas bolhas de Taylor. A seção 4.2 aborda a aplicação do fluxo óptico como ferramenta de rastreamento e estimativa de velocidade em sistemas bifásicos. Na seção 4.3, são apresentados estudos nos quais o modelo YOLO foi utilizado em conjunto com o método de fluxo óptico. Finalmente, a seção 4.4 traz uma conclusão, exemplificando a aplicação desses métodos no caso de estudo desenvolvido nesta dissertação.

4.1 Métodos de Classificação e Detecção

No contexto dos escoamentos multifásicos, um dos temas mais recorrentes é a classificação dos diferentes regimes de escoamento. Kadish et al. (2022), ao analisar filmagens de escoamentos bifásicos verticais de CO_2 -água, propôs a utilização de uma Rede Neural Convolucional (RNC) associada a uma rede de memória de longo e curto prazo (LSTM). O objetivo foi desenvolver um método de alto desempenho totalmente baseado em imagens de câmeras para classificar o regime de escoamento vertical e criar uma ferramenta de estimativa da qualidade do vapor. Os resultados obtidos foram significativos, com o modelo apresentando uma acurácia de 92%, demonstrando a aplicabilidade e eficácia da abordagem proposta.

Contudo, esses processos de classificação dependem do contexto geral das imagens analisadas. Assim, para identificar objetos específicos nas imagens, o uso de detectores de objetos torna-se de grande relevância. A arquitetura YOLO, em suas diversas versões, tem apresentado resultados promissores na detecção em tempo real em diversas áreas. Por exemplo, em sistemas de vigilância de rodovias, Chen et al. (2023) reportaram uma acurácia de 90% utilizando um banco de dados online, superando outros algoritmos na contagem e validação de veículos. Além disso, na avaliação de defeitos em chips eletrônicos em tempo real, Wang et al. (2023) obtiveram um valor de precisão média (MPM) de 87,4%. Na área médica, Li et al. (2023) utilizaram o algoritmo para detectar tricomas vaginais, alcançando um MPM de 92%. Diversas outras aplicações do algoritmo podem ser encontradas em Rane (2023), que descrevem várias implementações na Indústria 4.0 e 5.0, destacando o modelo como estado-da-arte.

No contexto de escoamentos multifásicos, alguns autores exploraram a aplicação do detector YOLO. Entretanto, até o momento da redação desta dissertação, a utilização desse modelo tem sido restrita à detecção de bolhas comuns, sem incluir as bolhas de

Taylor. A seguir, discutirá-se o potencial do uso do YOLO para a detecção de bolhas alongadas em escoamentos desse tipo. Zhou et al. (2024) demonstraram a aplicação de um detector de bolhas de vapor no funcionamento de um trocador de calor, alcançando um treinamento com MPM de 89%. Já Zhao et al. (2022), com um MPM de 96%, utilizaram um rastreamento baseado no filtro de Kalman para acompanhar o deslocamento isolado de cada bolha detectada ao longo dos quadros, com uma taxa máxima de processamento de 82 FPS. Torisaki e Miwa (2020) desenvolveram um modelo capaz de detectar todas as bolhas em escoamentos com frações de vazios menores que 10%, uma vez que valores superiores resultam em sobreposição das bolhas.

4.2 Método de Fluxo Ótico

A utilização de campos vetoriais de escoamento é fundamental para o estudo do comportamento das fases ao longo do processo. Um dos métodos amplamente empregados para essa finalidade é a Velocimetria por Imagem de Partículas (PIV, do inglês *Particle Image Velocimetry*), que é uma técnica experimental utilizada para medir campos de velocidade em escoamentos de fluido. Nesse método, partículas são dispersas no fluido em movimento, e uma sequência de imagens é capturada com o auxílio de uma câmera de alta velocidade. As partículas refletem a luz de uma fonte iluminadora, o que permite a obtenção de informações sobre a velocidade do fluido. Uma aplicação prática do PIV em bolhas de Taylor foi realizada por Hout et al. (2002), que utilizaram o método para estudar o comportamento do filme de líquido ao redor da bolha formada. Embora os resultados obtidos tenham sido interessantes, o método se apresenta como uma abordagem cara e complexa, exigindo equipamentos especializados, como emissores de luz, e um processo de análise intensivo.

Dessa forma, o método de fluxo ótico se configura como uma alternativa viável para a avaliação do fluxo vetorial de movimento, utilizando a variação de intensidade dos pixels ao longo dos quadros de uma sequência de imagens. A literatura sobre o uso desse método foca principalmente em aplicações de rastreamento de objetos. Por exemplo, Qimin et al. (2014) aplicaram o método, com o auxílio de uma câmera capaz de capturar 30 FPS, para determinar a velocidade de um veículo em movimento, utilizando dados de um sistema de posicionamento global (*Global Positioning System* - GPS) acoplado ao carro para validação. Em comparação com os dados de referência, o método apresentou um erro médio de velocidade de 0,121 m/s, demonstrando a aplicabilidade do método para a verificação da velocidade de veículos em estradas. Em outro estudo, Ding et al. (2021) utilizaram o método de fluxo ótico proposto por Lucas-Kanade, combinado com o algoritmo piramidal, para analisar o comportamento de bolhas de cavitação em um vaso sanguíneo durante um procedimento de trombólise por ultrassom. O experimento resultou na formação de um campo vetorial de movimentação das bolhas, permitindo o

monitoramento de seu movimento e sendo útil para a otimização de parâmetros visando a melhoria da trombólise por meio do aumento da cavitação.

Até o momento da redação desta dissertação, não foram encontrados estudos aplicando o método de fluxo óptico, isoladamente ou combinado com o modelo YOLO, para a análise de bolhas de Taylor. No entanto, diversos estudos exploraram o método em escoamentos bifásicos, nos quais estão presentes bolhas. Por exemplo, Devasagayam et al. (2022) aplicaram o método para investigar o comportamento de partículas em microfluídicos, utilizando filmagens e o método de fluxo óptico com detecção de cantos, conforme a abordagem de Shi-Tomasi. A análise do perfil de velocidade, comparando com dados da literatura e dados experimentais, indicou que o método obteve resultados confiáveis e eficientes do ponto de vista computacional para investigações de escoamentos no sistema avaliado. Já Seong et al. (2023), ao estudarem bolhas de vapor formadas em um trocador de calor, ampliaram a aplicação do método ao integrá-lo a uma Rede Neural Convolucional (RNC) para a detecção e rastreamento das bolhas ao longo dos quadros de um vídeo. Além de realizar a detecção e o rastreamento com precisão de 90%, o método também possibilitou, por meio da segmentação das bolhas medida pelos fluxos ópticos vetoriais, a avaliação da taxa de crescimento médio das bolhas, bem como da frequência de nucleação ao longo do tempo.

4.3 Integração de Detector com Método de Fluxo Óptico

Como apresentado nas seções anteriores, o YOLO e o método de fluxo óptico se destacam como ferramentas relevantes na área de visão computacional, com resultados satisfatórios em suas respectivas aplicações, especialmente em termos de acuracidade e tempo de processamento. Nesse contexto, investigar a aplicação conjunta dessas ferramentas para compreender como sua integração pode contribuir para a detecção e rastreamento de objetos configura-se como uma área de estudo relevante.

O uso do método de fluxo óptico em combinação com outras ferramentas de detecção já foi explorado em diversos estudos. Um exemplo é o trabalho de Seong et al. (2023), que, ao utilizar uma Rede Neural Convolucional (RNC) em conjunto com o método de fluxo óptico, não apenas rastrearam bolhas ao longo dos quadros, mas também extraíram seu comportamento estrutural ao longo do tempo.

No que se refere à combinação específica do YOLO com o método de fluxo óptico, diversos estudos propuseram diferentes integrações para aplicações em diversas áreas. Em Vu e Chung (2017), foi sugerido o uso de um sistema integrado para rastreamento em tempo real em sistemas de segurança. O foco do trabalho foi integrar o método de Lucas-Kanade (LK) com o método de Shi-Tomasi para rastrear objetos ao longo do tempo. A inovação do estudo consistiu na utilização do YOLO como detector inicial para identificar objetos de interesse nos quadros. Em resumo, o algoritmo verificava, através

do YOLO, a presença de objetos de interesse em um quadro; se confirmada sua presença, o método de Shi-Tomasi era aplicado apenas nas caixas delimitadoras formadas, e, em seguida, o método de fluxo óptico de Lucas-Kanade era utilizado para calcular os vetores de movimento. Essa estratégia permitiu reduzir a aplicação do método de fluxo óptico em toda a imagem, concentrando-se apenas nos objetos de interesse, o que resultou em um aumento na eficiência do processamento, alcançando uma taxa máxima de processamento de 145 FPS, com uma acuracidade mínima de 96,19%.

Em Liu et al. (2018), foi proposto um sistema de detecção e rastreamento de objetos que utiliza o fluxo óptico para calcular a velocidade dos objetos. O algoritmo desenvolvido avaliava cada quadro, detectando os objetos de interesse e calculando o fluxo óptico apenas nas regiões delimitadas, permitindo obter a velocidade dos objetos em tempo real. Com os vetores calculados, a velocidade média foi determinada pela proporção entre metros e pixels, além de ser realizada a correção ótica para ajustar a posição da câmera, dado que a movimentação não ocorria exclusivamente nas direções horizontal e vertical. Este método demonstrou ser eficaz na detecção da velocidade de objetos em tempo real.

No trabalho de Kaputa e Landy (2021), foi desenvolvido um algoritmo para construção de mapas geográficos em sistemas robóticos autônomos. O YOLO foi empregado para detectar objetos como pessoas e placas de sinalização, enquanto o método de fluxo óptico (OF) foi utilizado para desconsiderar objetos em movimento na construção do cenário global, uma vez que esses objetos, por não serem estáticos, não representam a configuração geográfica. O sistema alcançou uma acuracidade de 89,54%.

Utilizando o YOLO como detector e o método de fluxo óptico de Lucas-Kanade em pirâmides, Tian et al. (2023) desenvolveram uma técnica para detectar e avaliar o deslocamento de marcadores faciais com o objetivo de reconstruir o movimento em 3D. O YOLO foi utilizado para identificar os marcadores nos rostos dos atores, e cada marcador detectado foi analisado com base nos vetores de deslocamento de pixel calculados pelo fluxo óptico. Esses vetores foram fundamentais para a avaliação do movimento facial e para a reconstrução 3D. Embora seja necessário um maior número de experimentos para aumentar a confiabilidade do detector e do processo, os resultados obtidos para o grupo experimental estudado indicaram um desempenho superior ao de rastreadores de última geração, como SiamDW (ZHANG; PENG, 2019) e ECO (DANELLJAN et al., 2017). As métricas de erro quadrático médio e área sob a curva confirmaram melhorias na precisão e estabilidade dessa abordagem.

4.4 Conclusão

A revisão da literatura realizada neste trabalho revela um panorama promissor para a aplicação dos métodos de visão computacional, em especial o YOLO e o método de fluxo óptico, na detecção e análise de objetos em imagens. Embora estudos anteriores

tenham demonstrado um bom desempenho do YOLO na detecção de objetos em áreas correlatas aos estudos de escoamentos, não foi encontrado na literatura nenhum estudo que tenha aplicado essas técnicas especificamente em bolhas alongadas. Da mesma forma, a extração de vetores de velocidade para bolhas alongadas utilizando o método de fluxo óptico também não foi amplamente abordada.

A combinação entre os métodos de YOLO e fluxo óptico se apresenta como uma abordagem híbrida promissora, a ser investigada no contexto de escoamentos com bolhas alongadas. Esta integração não só permite a detecção precisa e em tempo real das bolhas, mas também facilita a medição da velocidade e o acompanhamento do comportamento dinâmico dessas bolhas. O YOLO, ao atuar como detector, é capaz de identificar e delimitar as bolhas em filmagens, e, dentro dessas regiões delimitadas, é possível avaliar com mais precisão a velocidade apenas do objeto de interesse. Dessa forma, além da detecção acurada das bolhas, é possível concentrar a análise na região de interesse, evitando que o comportamento local seja ofuscado por informações provenientes do contexto global. Essa abordagem favorece uma análise detalhada do movimento das bolhas, proporcionando uma maior compreensão dos processos multifásicos verticais. Portanto, a combinação dessas técnicas será investigada ao longo deste trabalho para analisar o movimento das bolhas de Taylor em escoamentos com fluido contínuo.

5 Metodologia

5.1 Montagem Experimental

A montagem experimental, realizada no *Flow and Research Laboratory* (Flow&Rs) do Centro de Estudos de Energia e Petróleo (CEPETRO) da Universidade Estadual de Campinas (Unicamp), e ilustrada na Figura 15, foi concebida com o intuito de permitir a condução de experimentos controlados. O dispositivo principal consiste em um tubo de vidro transparente com diâmetro interno de 26,35 mm e diâmetro externo de 29,90 mm, cujas medições foram realizadas com um paquímetro de precisão de 0,05 mm. O tubo, com comprimento total de 2 metros, é posicionado verticalmente durante os experimentos.

Para controlar o escoamento do ar dentro do tubo preenchido com água, o sistema conta com uma válvula manual, localizada na parte inferior, e uma válvula automática próxima ao topo do tubo. A válvula manual permite o controle do fluxo de ar por meio da abertura ou fechamento, com a entrada de ar sendo facilitada pela diferença de pressão, sem o uso de bomba de alimentação. Para garantir a estabilidade estrutural, duas hastes de aço suportam firmemente o dispositivo experimental.

Uma câmera fotográfica, modelo *Nikon D810*, foi montada sobre um tripé posicionado horizontalmente, a uma distância de aproximadamente 1 metro do tubo. A câmera foi configurada com uma abertura de $f/8$, e a taxa de captura de imagens foi ajustada para 59,94 quadros por segundo, com resolução de 1080p. A iluminação foi ajustada para garantir a constância da intensidade luminosa, sendo uma fonte de luz posicionada paralelamente à tubulação e oposta à câmera. Um plano de fundo de papel branco foi colocado entre a fonte de luz e o tubo, permitindo a dispersão uniforme da luz e assegurando uma iluminação constante durante a captura das imagens.

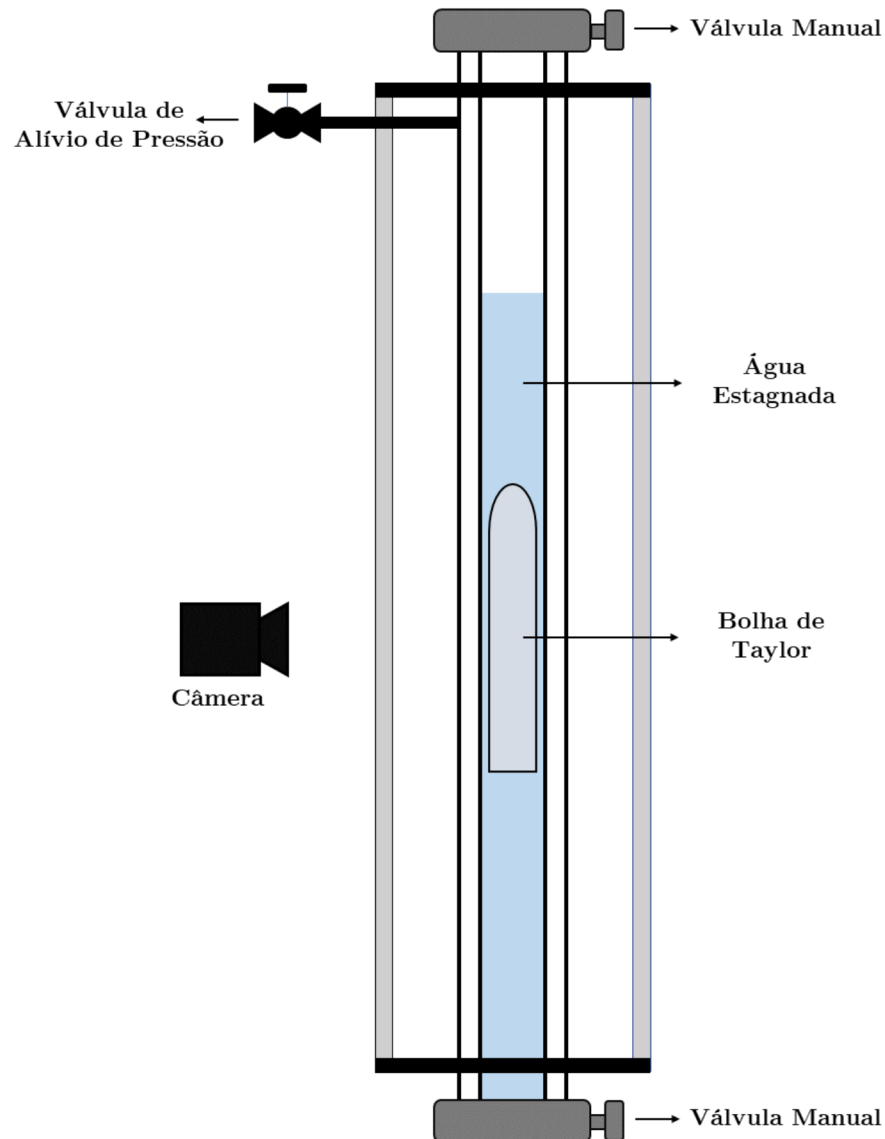
Além disso, foram utilizados dois refletores adicionais para fornecer uma iluminação difusa, essencial para garantir a qualidade das imagens capturadas, permitindo uma visualização nítida das estruturas no interior do tubo.

Este arranjo experimental foi projetado para proporcionar condições ideais de observação e análise do fluxo de ar, garantindo controle sobre as variáveis de iluminação e captura de imagens, e, consequentemente, possibilitando a obtenção de resultados mais consistentes para os estudos realizados.

5.2 Procedimento Experimental

O experimento tem como objetivo a coleta de dados sobre as bolhas de ar alongadas em meio líquido, que serão posteriormente utilizadas para o desenvolvimento de um algoritmo de detecção baseado em métodos de fluxo óptico. As bolhas foram registradas em imagens dentro de um tubo vertical preenchido com água, o que possibilitou o estudo de suas características, como forma, tamanho e movimento, informações essenciais para

Figura 15: Esquema experimental para a coleta de dados em líquido estagnado.



o aprimoramento de um sistema automatizado de detecção.

O procedimento experimental foi organizado em etapas, conforme descrito a seguir:

- **Preenchimento da Tubulação:** A tubulação foi preenchida com água até a proximidade da borda, a fim de evitar transbordamentos durante o escoamento do ar.
- **Inicialização do Sistema:** Com as válvulas fechadas, o ar contido no reservatório foi liberado, iniciando o processo de formação das bolhas. Simultaneamente, a câmera foi ativada para registrar os dados visuais.
- **Abertura e Fechamento da Válvula:** A válvula manual foi operada para permitir a entrada do ar, gerando uma bolha de Taylor que se estendia ao longo do

comprimento do tubo.

- **Armazenamento dos Vídeos:** A captura de imagens foi interrompida após a formação das bolhas, e os vídeos obtidos foram armazenados para análise posterior.

Esse procedimento foi repetido para avaliar diversas bolhas de Taylor, priorizando diferentes comprimentos de bolha sempre que possível. Foram coletadas 14 amostras para o estudo em água, sendo 6 selecionadas aleatoriamente para o treinamento do detector de objetos, e as 8 restantes para o algoritmo de detecção, integrado ao método de fluxo óptico.

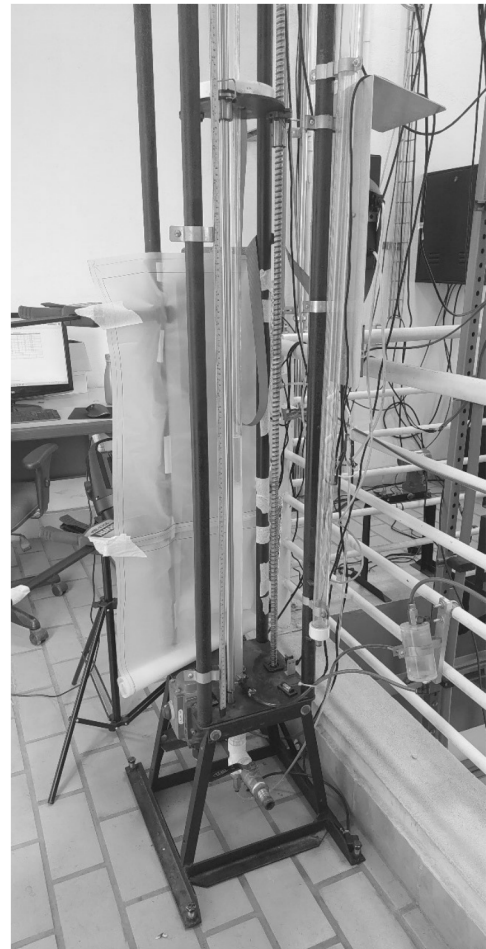
Os dados coletados foram analisados de maneira comparativa com os valores encontrados na literatura, utilizando equações e constantes específicas para validação.

Figura 16: Montagem experimental.

(a) Sistema de iluminação com luz de intensidade contínua.



(b) Tubulação de vidro onde foi realizada a visualização da bolha alongada.



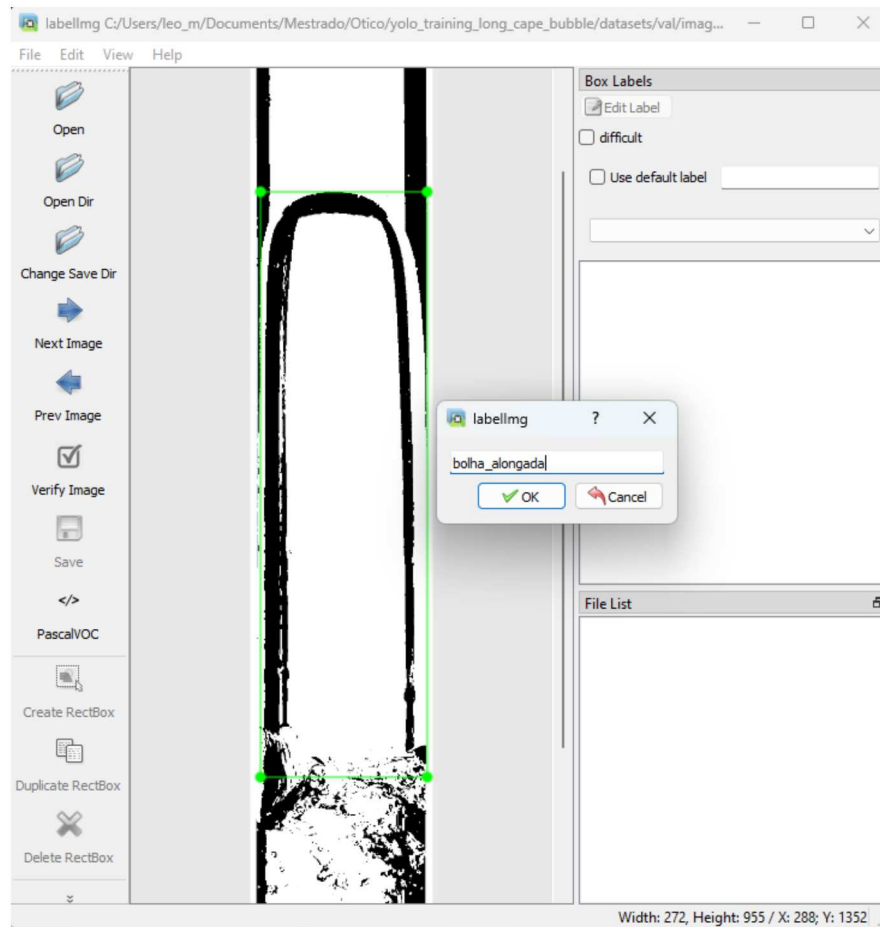
5.3 Detector de Objetos

Para utilizar o algoritmo de fluxo óptico, foi necessário criar um detector de imagens capaz de identificar e classificar bolhas alongadas em cada quadro. Para construir esse detector, foi utilizada a biblioteca *Ultralytics*, que implementa a arquitetura do YOLOv8. Como os objetos de interesse não estavam disponíveis nos modelos pré-treinados, foi necessário treinar o modelo para reconhecer as classes específicas. O processo incluiu as seguintes etapas:

- a) **Coleta e tratamento de dados:** Analisaram-se os vídeos coletados em condições similares para identificar bolhas alongadas. Os 14 vídeos destinados à etapa de detecção foram divididos aleatoriamente em dois grupos: um com 8 vídeos e outro com 6. A partir dos 8 vídeos, foram obtidas pelo menos 500 amostras de imagens de bolhas alongadas, capturando quadros de todos os vídeos.

Para criar um modelo de detecção mais abrangente, as imagens foram binarizadas, ou seja, convertidas para uma escala de intensidades de 0 e 255. Considerando que as condições de iluminação são constantes, utilizou-se 10 imagens aleatórias com bolhas de Taylor para determinar um valor de um limiar. Nesse processo, pixels com intensidades acima desse valor foram definidos como 255 e os demais como 0, garantindo qualitativamente que a bolha fosse claramente delimitada, conforme ilustrado na Figura 17. Esse método foi aplicado então para todas as imagens, possibilitando uma detecção mais rápida e precisa das bolhas.
- b) **Identificação dos objetos:** Utiliza-se a ferramenta *"Label Image"*, que permite anotar imagens com caixas de detecção delimitando os objetos de interesse e classificando-os como bolhas alongadas;
- c) **Separação de dados:** As imagens coletadas foram divididas aleatoriamente em três conjuntos representando, respectivamente, 70%, 15% e 15% do total das amostras, para treinamento, teste e validação;
- d) **Treinamento do detector:** Com os dados divididos, foram avaliados diferentes hiperparâmetros para o treinamento, como número de épocas, modelo e quantidade de lotes (*batches*), visando avaliar o desempenho do detector;
- e) **Avaliação do detector:** Utilizando os 6 vídeos restantes, não utilizados para o treinamento, o detector foi avaliado quanto ao tempo de processamento e à acurácia para cada vídeo. Esses desempenhos foram comparados para cada um dos cinco escalonamentos: *Nano* (n), *Small* (s), *Medium* (m), *Large* (l) e *Extra-large* (x).

Figura 17: Exemplo de utilização da ferramenta “Label Image” para anotar a imagem, delimitando e classificando uma bolha alongada em uma das amostras.



5.4 Algoritmo de Rastreamento por Fluxo Óptico

Após o ajuste e a validação do modelo de detecção, é possível identificar a presença de bolhas alongadas em cada quadro. Para isso, é necessário aplicar o filtro limiar previamente avaliado. O detector gera um escore de precisão relacionado ao objeto detectado, o qual indica a confiabilidade da classificação do modelo para cada detecção realizada. Além disso, o modelo delimita a região onde o objeto detectado está localizado. O algoritmo de fluxo óptico emprega a comparação das intensidades de pixels para rastrear e medir o deslocamento de objetos entre dois quadros consecutivos, com foco na região da imagem onde o objeto de interesse foi detectado, em comparação com a mesma região no quadro anterior.

Assim, para que o processo de rastreamento seja eficiente, o processamento da imagem segue as etapas descritas abaixo:

a) Tratamento de Imagem

Com o objetivo de aprimorar a qualidade das imagens e melhorar o desempenho do algoritmo de detecção, os quadros a serem analisados pelo método de fluxo óptico

passam por um tratamento específico. O processo foi implementado por meio de um filtro que realiza as seguintes etapas:

- i) **Normalização:** Inicialmente, a imagem é normalizada para garantir que todos os valores de pixel estejam dentro de uma faixa padrão. Esse processo transforma os valores de pixel para um intervalo de 0 a 255. A normalização é realizada subtraindo-se o valor mínimo de pixel da imagem e, em seguida, dividindo-se pela diferença entre o valor máximo e o valor mínimo. O resultado é escalado para a faixa de 0 a 255;
- ii) **Delimitação:** Após a normalização, é aplicada uma delimitação baseada em um valor de limiar. Pixels com valores abaixo desse limiar são definidos como 0. Esta etapa tem como objetivo eliminar ruídos ou detalhes irrelevantes de menor intensidade, que poderiam interferir na detecção precisa das bolhas;
- iii) **Filtro Mediano:** Em seguida, é aplicado um filtro mediano para reduzir o ruído e suavizar a imagem. O filtro mediano substitui cada pixel pelo valor mediano dos pixels em sua vizinhança, definida pelo tamanho do filtro. Este processo auxilia na preservação das bordas das bolhas, ao mesmo tempo em que elimina ruídos pontuais;
- iv) **Recorte da Região:** Por fim, utilizando a posição da caixa delimitadora obtida pela detecção da bolha mais recente, é recortada uma porção da imagem previamente avaliada. Esse recorte visa apresentar apenas as extremidades das bolhas de interesse, pois é nessa região que se deseja obter a velocidade do objeto.

Essas etapas combinadas resultam em imagens mais nítidas, definidas e delimitadas, facilitando o rastreamento das bolhas pelo algoritmo de fluxo óptico.

b) Método Lucas-Kanade

Dentre os métodos de rastreamento por fluxo óptico, o método local proposto por Lucas-Kanade é adequado quando atendidos os seguintes critérios:

- i) Os quadros comparados têm um intervalo de tempo curto entre as capturas, o que assegura que o objeto não tenha se deslocado abruptamente. Essa condição é comum em filmagens realizadas com câmeras de alta taxa de captura ou para objetos de baixa velocidade;
- ii) Os objetos de interesse apresentam contornos bem definidos em uma região delimitada (extremidade da bolha), o que se reflete em gradientes de intensidade notáveis nas imagens.

Considerando que as filmagens foram capturadas a 60 quadros por segundo, com um intervalo curto entre as imagens, e que os objetos de interesse possuem contornos bem definidos, optou-se pela aplicação do método Lucas-Kanade para o rastreamento utilizando o algoritmo de fluxo óptico.

Para a implementação, foi utilizada a biblioteca OpenCV, que oferece uma variedade de algoritmos para detecção e rastreamento de objetos, além de outras ferramentas. Um dos algoritmos disponíveis no OpenCV é o rastreamento por fluxo óptico LK, que possibilita o uso de pirâmides de níveis. Além disso, foi utilizado o detector de cantos baseado no método Shi-Tomasi para avaliar os pontos iniciais de rastreamento. Esse algoritmo é apropriado para a quantificação do fluxo óptico.

Como entrada, as duas imagens consecutivas, previamente tratadas, são inseridas juntamente com os parâmetros que influenciam o desempenho do algoritmo. Entre esses parâmetros, destacam-se:

- i) **maxCorners**: o número máximo de pixels a serem selecionados para rastreamento. Caso sejam detectados mais pixels, apenas os de maior intensidade serão selecionados;
- ii) **qualityLevel**: o valor mínimo para que um pixel seja selecionado, geralmente inferior a 1. Por exemplo, se o pixel de maior qualidade tiver um valor de 1000 e o parâmetro **qualityLevel** for 0.1, apenas os pixels com qualidade maior ou igual a 100 serão selecionados;
- iii) **minDistance**: a menor distância euclidiana entre os pixels selecionados;
- iv) **blockSize**: o tamanho do bloco de pixels ao redor do pixel avaliado para identificar contornos;
- v) **maxLevel**: o número máximo de níveis de pirâmide que o método utilizará. Definir este parâmetro como 0 significa que apenas a imagem original será utilizada;
- vi) **winSize**: o tamanho da janela média usada pelo algoritmo. Janelas maiores aumentam a robustez do resultado, mas exigem maior capacidade de processamento.

Esses parâmetros devem ser ajustados de acordo com as características do conjunto de dados, como o tamanho dos quadros, a luminosidade das filmagens e a intensidade dos pixels. Assim, a avaliação desses parâmetros se torna necessária para cada tipo de escoamento estudado, garantindo que todos os quadros analisados apresentem condições de iluminação semelhantes e possam ser processados com os mesmos critérios.

5.5 Tratamento Matemático dos Resultados

Ao integrar o detector e o rastreador, várias informações podem ser extraídas do algoritmo para cada quadro. O detector fornece a caixa delimitadora do objeto de interesse, permitindo determinar suas dimensões, localização, classe e a porcentagem de certeza associada à detecção. Por outro lado, o método de fluxo óptico gera vetores de deslocamento para todos os pixels detectados em movimento.

A partir do detector, é possível calcular informações como a velocidade da bolha de Taylor ($v_{BT, \text{detec}}$) e seu comprimento ($L_{BT, \text{detec}}$). A velocidade obtida pelo detector pode ser calculada utilizando a distância como a diferença entre as posições da ponta da bolha no eixo y ($p_{y, \text{primeiro quadro}}$) no primeiro quadro, em comparação com a posição da ponta no quadro seguinte ($p_{y, \text{segundo quadro}}$). O tempo é determinado pela diferença entre o número do quadro atual e o número do quadro seguinte, multiplicada pelo intervalo de captura de tempo entre dois quadros consecutivos. Esse intervalo é expresso em segundos e corresponde ao inverso da taxa de quadros por segundo (FPS) da câmera de captura.

$$t_{\text{captura}} = \frac{1}{\text{FPS}} \quad (31)$$

$$v_{BT, \text{detec}} = \frac{p_{y, \text{segundo quadro}} - p_{y, \text{primeiro quadro}}}{(n_{\text{segundo quadro}} - n_{\text{primeiro quadro}}) \cdot t_{\text{captura}}} \quad (32)$$

Esses valores estão na escala de pixels, assim como ($L_{BT, \text{detec}}$), que é obtido diretamente pelo detector. Para converter as distâncias de pixels para metros, é necessária uma conversão de unidades. A conversão da distância em metros para pixels pode ser realizada medindo o diâmetro externo do tubo de acrílico no visor, em metros, e fazendo a mesma medição em pixels em um dos quadros de referência. A razão para utilizar o diâmetro externo do tubo, em vez do diâmetro interno, deve-se a um efeito de distorção provocado pela refração da luz nas interfaces do vidro. Embora teoricamente o diâmetro interno seja o valor mais adequado, uma vez que ele representa diretamente o espaço ocupado pela bolha, a distorção causada pela refração na porção interna do vidro dificulta a delimitação precisa da borda interna nas imagens. Dessa forma, o diâmetro externo é preferido, pois a medição pode ser feita de forma mais confiável, com menor interferência de efeitos de refração.

$$\text{prop}_{\text{m/pixel}} = \frac{\text{diâmetro em metros}}{\text{diâmetro em pixels}} \quad (33)$$

Para o método de fluxo óptico, o algoritmo gera uma matriz de vetores que representam o deslocamento da intensidade de cada pixel entre dois quadros consecutivos ($d_{\text{pixel, unit}}$). No entanto, assim como no detector, o valor do deslocamento é expresso em pixels, sendo necessária a aplicação da conversão apresentada na Equação 33. Assim, para

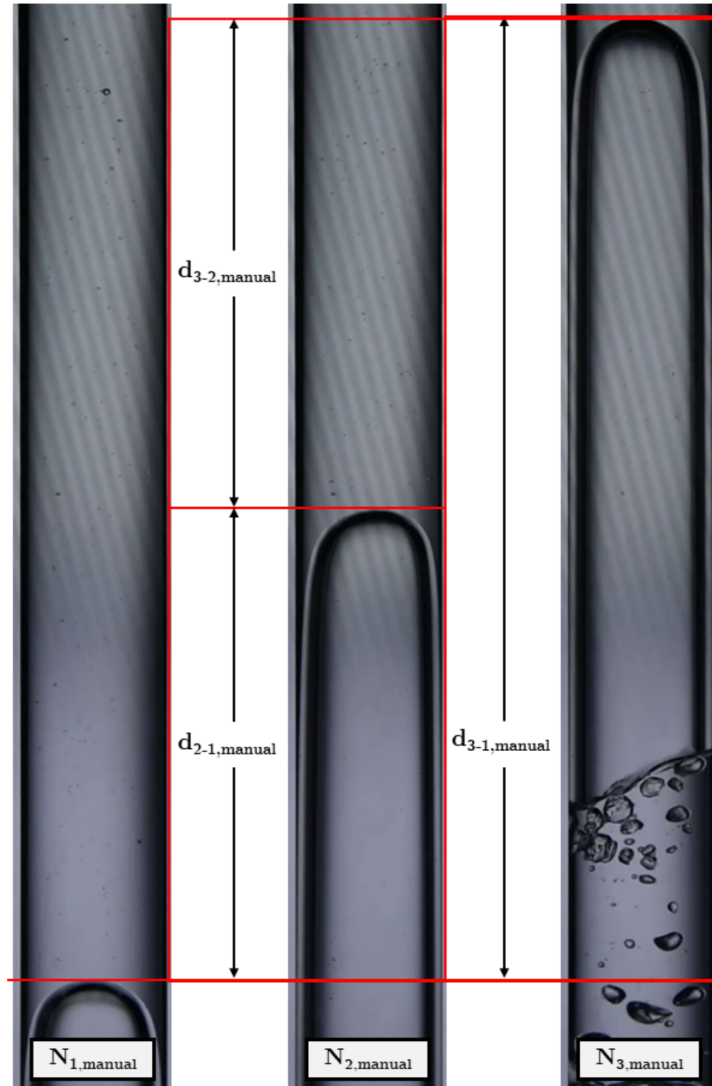
cada pixel analisado, a velocidade $v_{\text{pixel, unit}}$ em metros por segundo é dada por:

$$v_{BT,\text{óptico}} = \frac{d_{\text{pixel, unit}} \cdot \text{prop}_{\text{m/pixel}}}{(n_{\text{segundo quadro}} - n_{\text{primeiro quadro}}) \cdot t_{\text{captura}}} \quad (34)$$

A velocidade da bolha medida pelo método de fluxo óptico ($v_{BT,\text{óptico}}$) é então obtida pela média das velocidades dos pixels na região da ponta da bolha.

Além dos métodos automáticos, foi realizado um cálculo manual da velocidade, o qual é ilustrado esquematicamente na Figura 18. Essa avaliação consistiu na seleção de três quadros distintos para cada bolha apresentada nos vídeos: no primeiro quadro, a ponta da bolha foi localizada no início do vídeo; no segundo quadro, no meio; e no terceiro quadro, no final do vídeo. Para cada um desses quadros, foi registrado o número do quadro ($N_{i,\text{manual}}$) dentro do total de quadros do vídeo.

Figura 18: Método esquemático utilizado para o cálculo manual da velocidade.



Nos três quadros, a posição da ponta das bolhas foi selecionada por um especialista

humano, permitindo determinar a distância vertical entre os pontos, resultando em uma distância $d_{i,\text{manual}}$ em pixels entre as pontas. Com esses dados, a velocidade entre um quadro i e um quadro j foi calculada pela seguinte equação:

$$v_{j-i,\text{manual}} = \frac{d_{j-i,\text{manual}}}{(N_{j,\text{manual}} - N_{i,\text{manual}}) \cdot t_{\text{captura}}} \cdot \text{prop}_{\text{m/pixel}} \quad (35)$$

A velocidade da bolha obtida pelo método manual ($v_{BT,\text{manual}}$) é, então, calculada pela média aritmética das velocidades associadas às distâncias d_{2-1} , d_{3-1} e d_{3-2} . O número do quadro (N) corresponde ao índice do quadro avaliado dentro do total de quadros do vídeo.

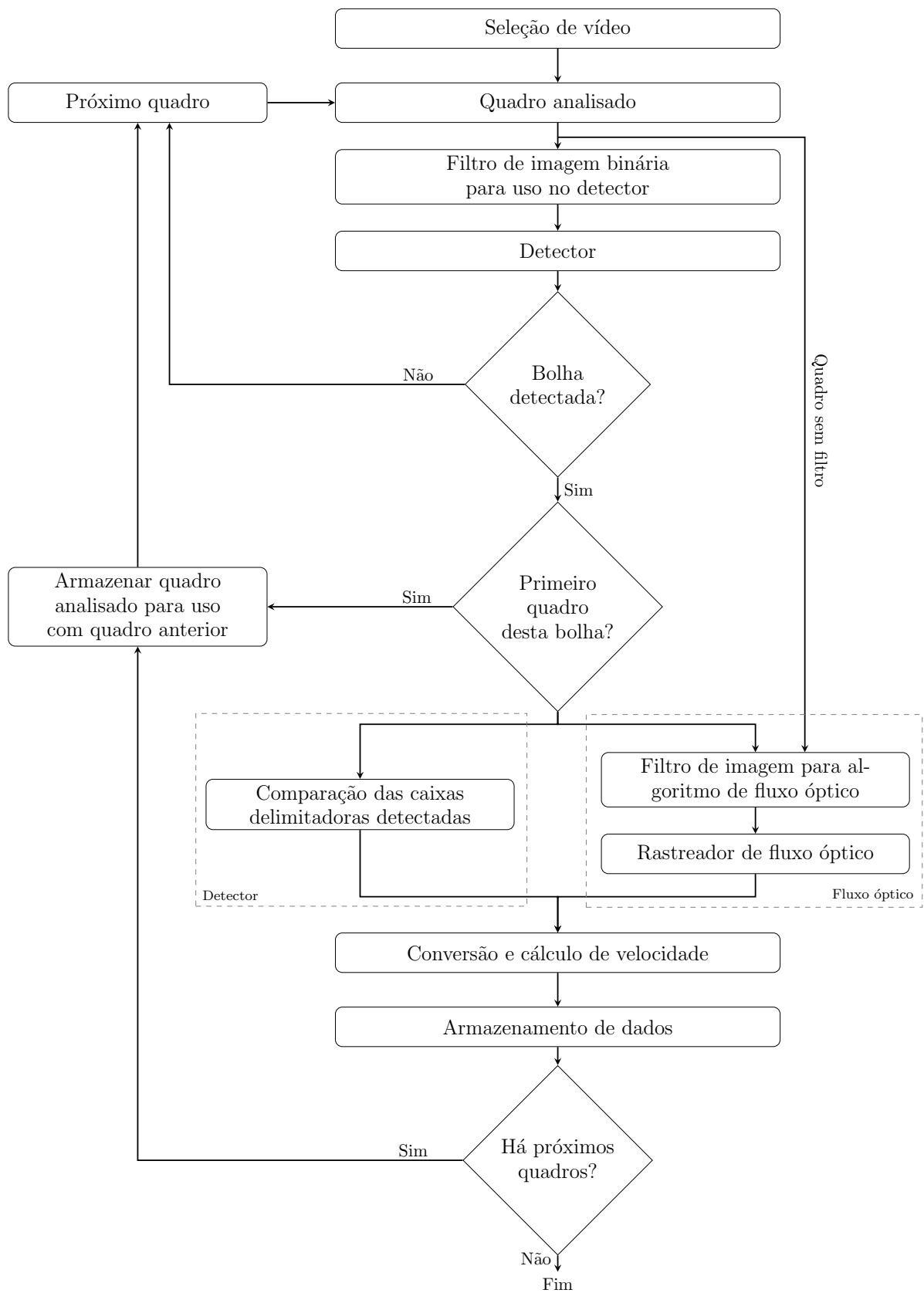
$$v_{BT,\text{manual}} = \frac{v_{2-1,\text{manual}} + v_{3-1,\text{manual}} + v_{3-2,\text{manual}}}{3} \quad (36)$$

Com as velocidades da bolha de Taylor calculadas, também será possível avaliar a constante C_t , que pode ser determinada isolando-a na equação 15, para as velocidades $v_{BT,\text{teórico}}$, $v_{BT,\text{óptico}}$ e $v_{BT,\text{manual}}$.

5.6 Funcionamento Global do Algoritmo

A Figura 19 apresenta o funcionamento esquemático do algoritmo. Primeiramente, é dado como entrada o vídeo de estudo, que é selecionado, sendo avaliado quadro a quadro. O quadro passa primeiramente por um filtro de imagem binário, isso se deve ao fato do treinamento do detector ter sido totalmente nessa escala, permitindo uma padronização no momento de detecção. Em seguida, para o detector, a imagem é avaliada pelo modelo previamente treinado, resultando nas caixas de detecção encontradas. Se estas possuírem um percentual de certeza maior que 0,75, toma-se que uma bolha alongada foi detectada, caso contrário, avalia-se o próximo quadro. Detectada uma bolha, há então a avaliação, através de um contador, se o quadro avaliado é o primeiro quadro em que ela está sendo detectada; se sim, esse quadro é armazenado e o próximo é avaliado. Caso a esta bolha já tiver sido detectada, este quadro e o anterior armazenado passarão por um filtro e corte de imagem. O filtro terá como objetivo realçar os pixels dos contornos de maneira a facilitar a utilização do método do fluxo óptico, já o corte terá como objetivo avaliar apenas a parte central do quadro em que as bolhas foram detectadas, pois dessa maneira estará se avaliando a velocidade na ponta da bolha, correspondendo à velocidade translacional $v_{BT,\text{óptica}}$. As devidas operações matemáticas apresentadas na Seção 5.5 são realizadas e os valores $v_{BT,\text{óptico}}$, $v_{BT,\text{detec}}$, $L_{BT,\text{detec}}$ entre outras propriedades de interesse são armazenados. Caso não haja mais quadros no vídeo, o algoritmo será encerrado.

Figura 19: Fluxograma lógico do algoritmo construído.



5.7 Descrição do Hardware e Versões de Software

O hardware adotado compreende uma placa-mãe ASUS TUF Gaming Z690-PLUS D4, um disco SSD Kingston SNV2S/1000G M.2 de 1 TB, quatro módulos de memória RAM Kingston DDR4-3200 MHz de 32 GB cada, uma placa de vídeo NVIDIA GeForce RTX 4060 Ti com 16 GB de memória, e um processador Intel Core i7-13700 KF com 16 núcleos operando a 3,4 GHz. O sistema operacional utilizado é baseado em Linux, na distribuição Ubuntu 22.04.2 LTS com arquitetura X86 64.

Para o processamento de dados, o ambiente de desenvolvimento utilizou a linguagem Python na versão 3.10.11. Diversas bibliotecas foram empregadas para desenvolver e avaliar o desempenho dos modelos de aprendizado profundo, tais como OpenCV 4.9.0, Ultralytics 8.1.30 e Numpy 1.23.5. Para visualização dos resultados, foram utilizadas Pandas 1.5.3, Matplotlib 3.7.1, Seaborn 0.12.0, Scikit-image 0.22.0 e Sklearn 1.2.2.

6 Resultados

6.1 Treinamento do Detector de Objetos

Como discutido na Seção 3.2.4, a arquitetura YOLO permite diferentes escalonamentos, que ampliam o tamanho das camadas convolucionais. Dependendo dos dados empregados e da capacidade computacional disponível, esses escalonamentos podem influenciar a média da precisão média (mAP) do modelo, resultando em uma detecção mais eficaz dos objetos de interesse. Para explorar essa questão, foram realizados treinamentos de cinco detectores nos escalonamentos *Nano* (n), *Small* (s), *Medium* (m), *Large* (l) e *Extra-large* (x), utilizando um tamanho de lote (*batch size*) igual a 8 e um número de épocas igual a 100.

Tabela 2: Resultados dos detectores para diferentes escalonamentos.

Modelo	mAP IoU(50-95)	Avaliação Média dos Vídeos			
		Tempo Real (s)	Tempo Algoritmo (s)	Velocidade (FPS)	Acurácia
<i>Nano</i>	84,5%	1,85	0,81	137	87%
<i>Small</i>	84,3%		0,82	137	88%
<i>Medium</i>	84,2%		0,93	120	84%
<i>Large</i>	84,6%		1,04	107	83%
<i>Extra-large</i>	83,4%		1,47	76	77%

A Tabela 2 apresenta os resultados das médias das precisões médias dos treinamentos, considerando os valores de região sobre união (IoU) entre 0,5 e 0,95. Além disso, mostra a avaliação média de cinco vídeos distintos dos utilizados no treinamento, contendo bolhas de Taylor. O tempo real representa a média dos tempos dos vídeos originais, gravados a 60 quadros por segundo, enquanto o tempo do algoritmo refere-se à média do tempo necessário para processar cada vídeo. A velocidade, em quadros por segundo (FPS), é calculada pela média da razão entre a quantidade total de quadros no vídeo e o tempo de processamento do algoritmo, conforme a Equação 37.

$$V_{FPS,i} = \frac{\text{Quantidade total de quadros no vídeo}}{\text{Tempo de processamento do algoritmo em segundos}} \quad (37)$$

Por fim, a tabela apresenta a acurácia média dos vídeos. Para cada quadro em que uma bolha foi detectada, o algoritmo atribui um valor de confiança, indicando a probabilidade de acerto com base no treinamento. Para cada vídeo, a acurácia é obtida pela média desses valores de confiança, sendo o valor final da tabela a média entre os cinco vídeos avaliados.

Ao analisar os mAP para todos os modelos, observa-se valores muito próximos entre si, o que sugere que esses modelos apresentaram desempenho semelhante. Isso indica que o aumento do tamanho da arquitetura não resultou em melhorias significativas para os dados utilizados. Além disso, ao estender essa avaliação para a aplicação em vídeos diferentes dos utilizados no treinamento, nota-se que quanto maior o modelo, maior o tempo de processamento, o que reduz a velocidade do algoritmo. Além disso, para o maior modelo (*Extra-Large*), observa-se até mesmo uma redução na acurácia média das bolhas identificadas bem como em sua mAP , possivelmente devido a uma maior especificidade do treinamento. Este fenômeno pode ser explicado pelo aumento na capacidade de memorizar características específicas do conjunto de treinamento, resultando em um possível sobreajuste (overfitting), o que limita sua capacidade de generalizar para novas imagens.

A escolha do modelo deve considerar não apenas o desempenho, mas também fatores como consumo energético, escalabilidade e adaptação a diferentes contextos. Modelos menores consomem menos energia e são mais adequados para dispositivos com recursos limitados, enquanto modelos maiores demandam maior infraestrutura computacional. A disponibilidade de dados também é determinante, uma vez que conjuntos limitados favorecem modelos menores, enquanto conjuntos extensos possibilitam que modelos maiores aprendam representações mais complexas.

Com base nesses aspectos, o modelo *Small* apresentou o melhor desempenho geral, equilibrando precisão, velocidade e capacidade de generalização.

6.2 Comprimento das Bolhas de Taylor

Nesta seção, são apresentados os resultados obtidos na determinação do comprimento das bolhas de Taylor nos vídeos analisados, utilizando o modelo de detecção baseado na arquitetura YOLO. O estudo considerou exclusivamente bolhas completamente visíveis no campo de visão da filmagem, pois o método adotado apresenta limitações na análise de bolhas cujo comprimento excede a região visualizada. O detector fornece caixas delimitadoras para os objetos detectados, possibilitando a medição da altura dessas caixas, que representa o comprimento das bolhas em pixels. Esses valores foram convertidos para metros utilizando a proporção $\text{prop}_{\text{m/pixel}}$, conforme descrito na Seção 5.5.

A distribuição dos comprimentos das bolhas foi avaliada por meio de gráficos de dispersão do tipo violino. Os resultados foram analisados sob dois aspectos: (i) a dispersão dos dados de detecção ($\text{Detec}_{\text{distrib}}$), que representa a distribuição dos comprimentos detectados, e (ii) a mediana dos comprimentos de detecção ($\text{Detec}_{\text{med}}$), que corresponde à mediana desses valores. As análises consideraram um intervalo de confiança correspondente a um desvio padrão, como ilustrado na Figura 20. A figura também apresenta os comprimentos medidos manualmente, a partir da seleção de cinco quadros distintos de

cada vídeo, com medição da altura das bolhas em pixels e subsequente conversão para metros. A Tabela 3 compila os valores médios, desvios padrões, medianas e quartis coletados para cada bolha dos vídeos analisados, bem como os valores calculados pelo método manual.

Figura 20: Avaliação do comprimento das bolhas pelo método de detecção, utilizando dispersão em violino e média com um desvio padrão.

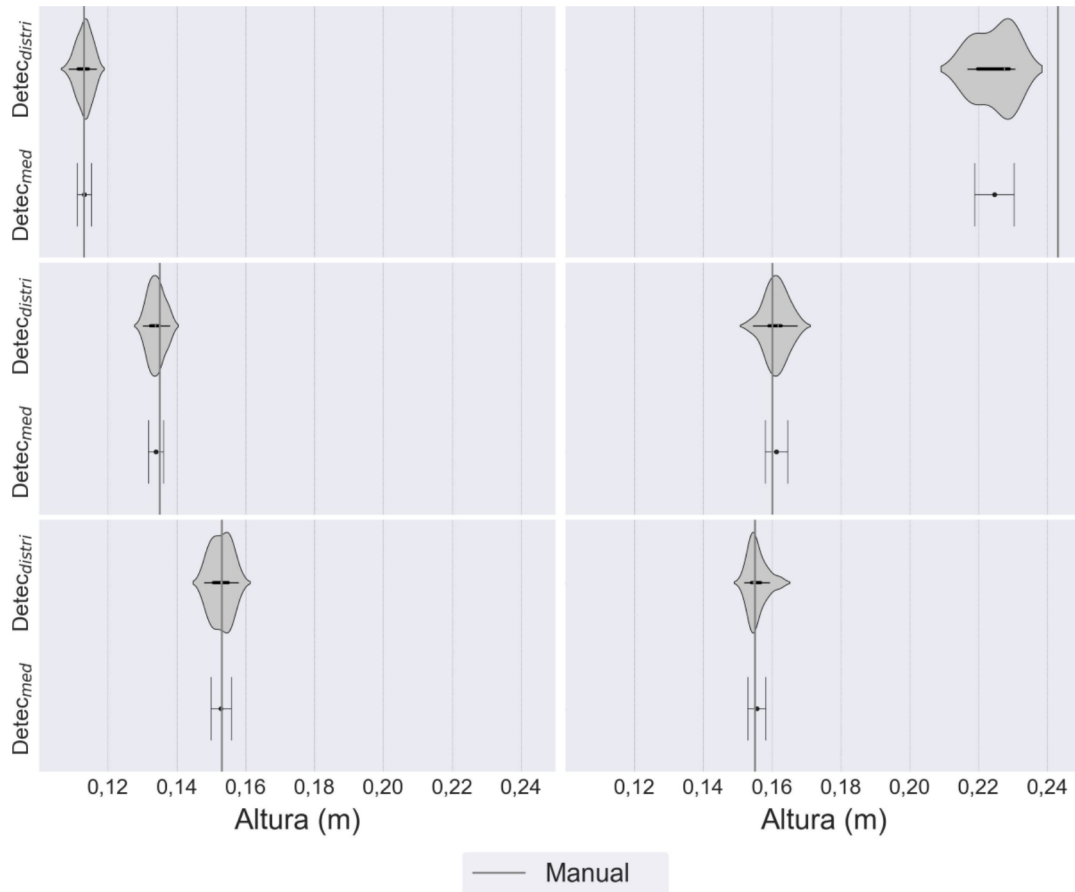


Tabela 3: Dados do comprimento das bolhas de Taylor.

Variáveis		Vídeos					
		1	2	3	4	5	6
$L_{BT,detec}(m)$	Média	0,113	0,225	0,134	0,161	0,153	0,156
	σ	0,002	0,005	0,002	0,003	0,003	0,003
	1Q	0,111	0,220	0,132	0,159	0,151	0,154
	Mediana	0,113	0,228	0,134	0,162	0,153	0,155
	3Q	0,114	0,229	0,135	0,163	0,155	0,157
$L_{BT,manual}(m)$		0,113	0,243	0,135	0,160	0,153	0,155

Ao avaliar as porcentagens significativas do desvio padrão em relação à média e a

distância interquartílica (entre o 1º e o 3º quartis) em relação à mediana, torna-se essencial analisar essas dispersões de forma percentual em relação ao valor avaliado. No método de detecção empregado, a dispersão máxima do comprimento das bolhas foi de 2,4% considerando o desvio padrão e 3,5% considerando a diferença interquartílica, indicando que o comprimento das bolhas obtido pelo método apresenta variações mínimas ao longo dos quadros analisados.

Para validar os dados obtidos, realizou-se uma comparação entre as médias e medianas calculadas pelo método de detecção com os valores coletados manualmente, apresentados na Tabela 4. Essa comparação evidenciou as diferenças percentuais entre os métodos, calculadas pela razão entre a diferença absoluta dos valores calculado e manual, dividida pelo valor manual.

Tabela 4: Comparação percentual entre os valores manuais com as médias e medianas dos comprimentos obtidos pelo método de detecção.

Variáveis		Vídeos					
		1	2	3	4	5	6
Manual (m):		0,113	0,243	0,135	0,160	0,153	0,155
$L_{BT,detec}$	Média	0,2%	7,4%	0,7%	1,1%	0,3%	0,5%
	Mediana	1,8%	9,5%	1,9%	0,2%	1,7%	0,5%

Nos vídeos 1 e de 3 a 6, observa-se que tanto as médias quanto as medianas dos comprimentos estão próximas dos valores medidos manualmente, com uma diferença máxima de 1,1% para as médias e 1,9% para as medianas. Ademais, os valores manuais situam-se dentro das faixas propostas, considerando a média com um desvio padrão e a diferença interquartílica. Esses resultados indicam que, para bolhas de comprimentos similares às avaliadas nesses vídeos, o método de detecção foi capaz de estimar o comprimento das bolhas de Taylor de forma satisfatória.

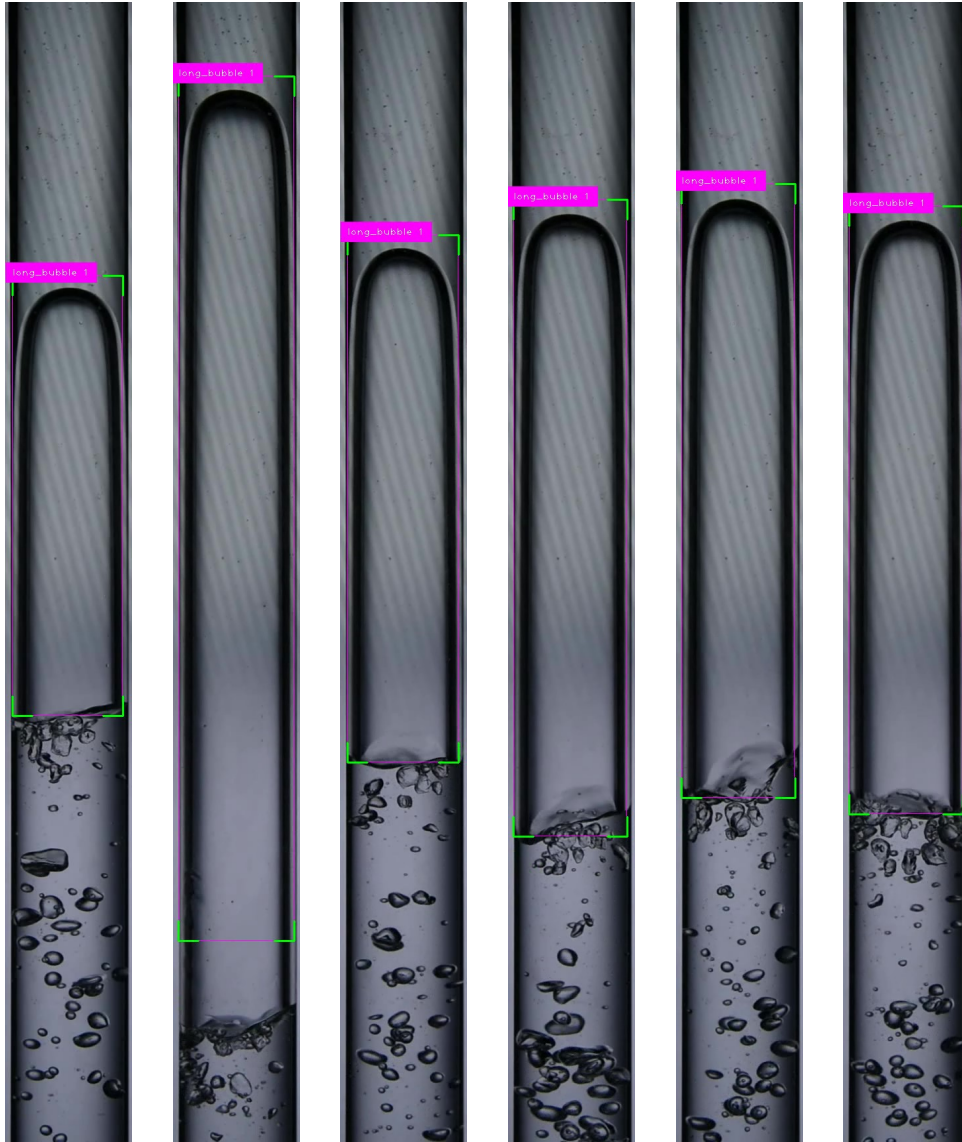
Entretanto, no vídeo 2, os resultados apresentam uma diferença percentual maior, com erros de 7,4% na média e 9,5% na mediana. Além disso, os valores calculados situam-se fora das faixas de desvio padrão e quartis. Qualitativamente, a Figura 21, que ilustra amostras da detecção em cada vídeo, evidencia que o método não conseguiu capturar a bolha em sua totalidade no vídeo 2, delimitando corretamente sua ponta, mas falhando em identificar adequadamente a extremidade final (calda) da bolha.

Essa limitação está associada ao fato de que o comprimento da bolha, aproximadamente 0,243 m, corresponde a cerca de 78% da extensão total do campo de visão (0,313 m). Assim, bolhas que se aproximam do limite do campo de visão enfrentam dificuldades de detecção, devido à limitação do modelo no treinamento para bolhas de maiores comprimentos. Esse problema não decorre de uma falta de precisão do método, mas sim de uma

restrição intrínseca da escala de medição.

Figura 21: Amostras da detecção nos vídeos avaliados.

(a) Vídeo 1 (b) Vídeo 2 (c) Vídeo 3 (d) Vídeo 4 (e) Vídeo 5 (f) Vídeo 6



Para aprimorar o método de detecção em futuros experimentos, recomenda-se fixar uma posição de referência na tubulação e implementar um detector capaz de identificar tanto a ponta quanto a calda da bolha de Taylor. Assim, o deslocamento da bolha pode ser medido a partir do quadro em que sua ponta cruza a posição até o quadro em que a calda finaliza a passagem. Essa abordagem mitigaria as limitações impostas pelo campo de visão restrito e permitiria a análise de bolhas de maiores comprimentos com maior robustez.

Os resultados obtidos confirmam que o método de detecção é eficaz para a maioria dos casos analisados, desde que as bolhas estejam completamente contidas no campo de visão da filmagem. Contudo, para bolhas de maior comprimento, como observado no

vídeo 2, a escala de medição representa uma limitação que deve ser abordada em trabalhos futuros.

6.3 Velocidade e Constante de Velocidade Terminal das Bolhas de Taylor

Nesta seção, serão apresentados os resultados de velocidade e a constante de velocidade terminal obtidos a partir do algoritmo utilizando um conjunto de 6 vídeos do experimento de escoamento de uma única bolha de Taylor em líquido estagnado, os quais não foram empregados em nenhuma etapa de treinamento. Essas médias não utilizadas no treinamento do detector permitem avaliar a funcionalidade e a generalização do algoritmo em novos dados.

6.3.1 Aplicação Delimitada do Método de Fluxo Óptico

Como apresentado na Seção 5.4, o método de fluxo óptico foi aplicado exclusivamente à região da imagem correspondente à ponta da bolha, uma vez que essa é a área de maior interesse para os objetivos do estudo. O método de fluxo óptico, por sua natureza, não realiza a distinção de objetos, baseando-se apenas nas variações de intensidade luminosa. Aplicar esse método em toda a imagem resultaria no cálculo de deslocamentos de pixels ao longo de toda a extensão da bolha, o que não é relevante para a análise pretendida.

Uma alternativa seria aplicar o método de fluxo óptico a toda a imagem e, posteriormente, filtrar os vetores de deslocamento na região de interesse. Contudo, tal abordagem aumentaria desnecessariamente o esforço computacional, já que o foco é exclusivamente a velocidade da ponta da bolha. Portanto, delimitar previamente a região de interesse permitiu reduzir significativamente o esforço computacional, sem comprometer a precisão dos resultados.

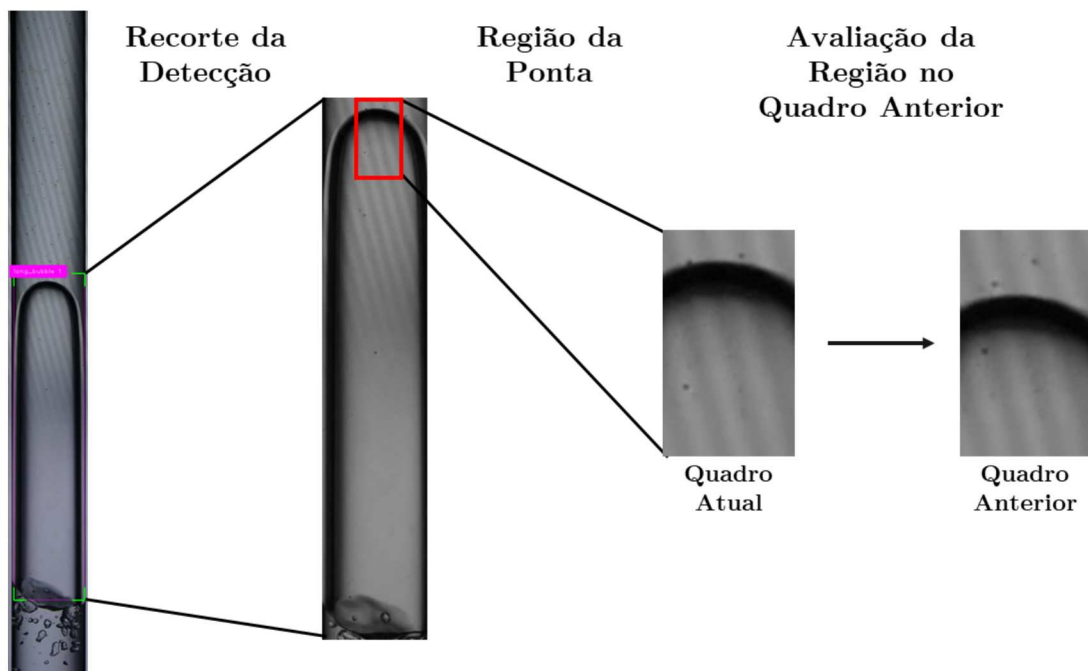
Para validar essa abordagem, foi realizada uma comparação utilizando os mesmos parâmetros para o vídeo ilustrado na Figura 22, aplicando o método com e sem a delimitação da região correspondente à ponta da bolha. A metodologia empregada para delimitar a região de interesse é detalhada a seguir.

A delimitação da região de interesse baseou-se na localização da ponta da bolha detectada pelo algoritmo YOLO. Considerou-se a porção central da seção transversal da detecção, dado que a ponta da bolha encontra-se no topo central em um escoamento vertical. Especificamente, selecionou-se uma faixa horizontal correspondente a 7,5% da largura total da imagem, centrada na detecção, e uma faixa vertical de 30% da altura da detecção, partindo do topo para baixo. Essa escolha, ilustrada na Figura 23, foi motivada pela necessidade de garantir que toda a curvatura da ponta da bolha fosse capturada. Para isso, adicionaram-se 10 pixels acima do início da bolha, assegurando a completa representação da curvatura na imagem.

Figura 22: Comparação entre a estimativa dos vetores de deslocamento pelo método de fluxo óptico Lucas-Kanade com e sem delimitação da região de interesse.



Figura 23: Delimitação da região para aplicação do método de fluxo óptico restrito à ponta da bolha.



Essa delimitação foi fundamental para eliminar a análise de deslocamento na região de esteira da bolha, onde o método de fluxo óptico pode apresentar maior dispersão de vetores devido às condições de baixa intensidade luminosa e movimentos mais difusos. Ao concentrar-se na ponta da bolha, a abordagem garantiu a análise de vetores de deslocamento em uma região com maior contraste e movimento bem definido, condições em que o método de fluxo óptico Lucas-Kanade opera de forma mais eficiente.

O método Lucas-Kanade calcula os vetores de deslocamento com base em pequenas janelas de pixels e presume uma variação linear da intensidade luminosa ao longo do tempo. Portanto, delimitar a região de aplicação evita a introdução de ruídos decorrentes de áreas com baixa correlação temporal, como a esteira, e assegura maior confiabilidade nos resultados obtidos. Essa estratégia ampliou o espaço amostral dos vetores de deslocamento úteis para a análise (região próxima a ponta da bolha), oferecendo mais dados relevantes para a caracterização do movimento da bolha.

Além disso, o tempo de processamento foi significativamente reduzido. Para o mesmo vídeo, o processamento de 40 quadros levou 19,2 segundos sem a delimitação e apenas 3 segundos com a delimitação. Essa diferença evidencia uma redução de 6,4 vezes no tempo de processamento, demonstrando a eficiência computacional da abordagem delimitada, sem comprometer a qualidade dos dados obtidos.

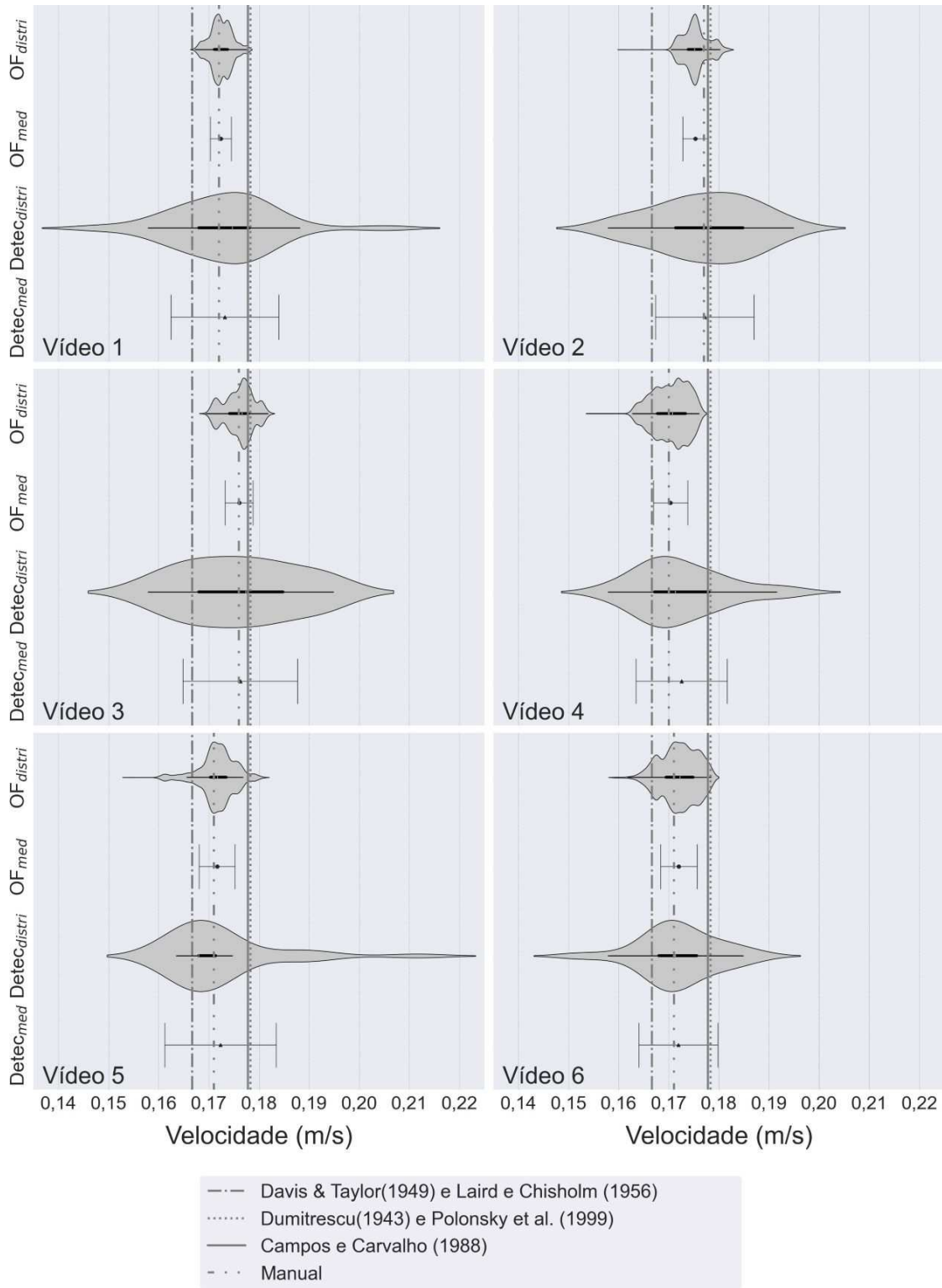
6.3.2 Avaliação da Dispersão dos Dados de Velocidade e Constante Terminal

Para cada vídeo, foi utilizada a dispersão do tipo violino para avaliar a distribuição das velocidades das bolhas. Os resultados foram analisados de duas formas: a dispersão dos dados de detecção ($\text{Detec}_{\text{distr}_i}$), que representa a distribuição das velocidades detectadas, e a mediana das velocidades de detecção ($\text{Detec}_{\text{med}}$). Além disso, avaliou-se a dispersão das velocidades calculadas pelo método de fluxo óptico integrado ao detector ($\text{OF}_{\text{distr}_i}$) e a mediana dessas velocidades (OF_{med}). As análises consideraram um erro equivalente a um desvio padrão, resultando nas Figuras 24 e 25. As figuras também apresentam as velocidades teóricas calculadas pela Equação 15, com os valores das constantes listados na Tabela 7, juntamente com as velocidades encontradas pelo método manual para cada bolha, conforme a Equação 36.

As Tabelas 5 e 6 apresentam os resultados médios coletados pelos métodos de detecção, fluxo óptico e manual para cada um dos vídeos de 1 a 6. A média e os erros calculados por um desvio padrão foram utilizados, uma vez que todos os dados coletados derivam da diferença entre dois quadros, resultando em uma média e desvio padrão dos mesmos.

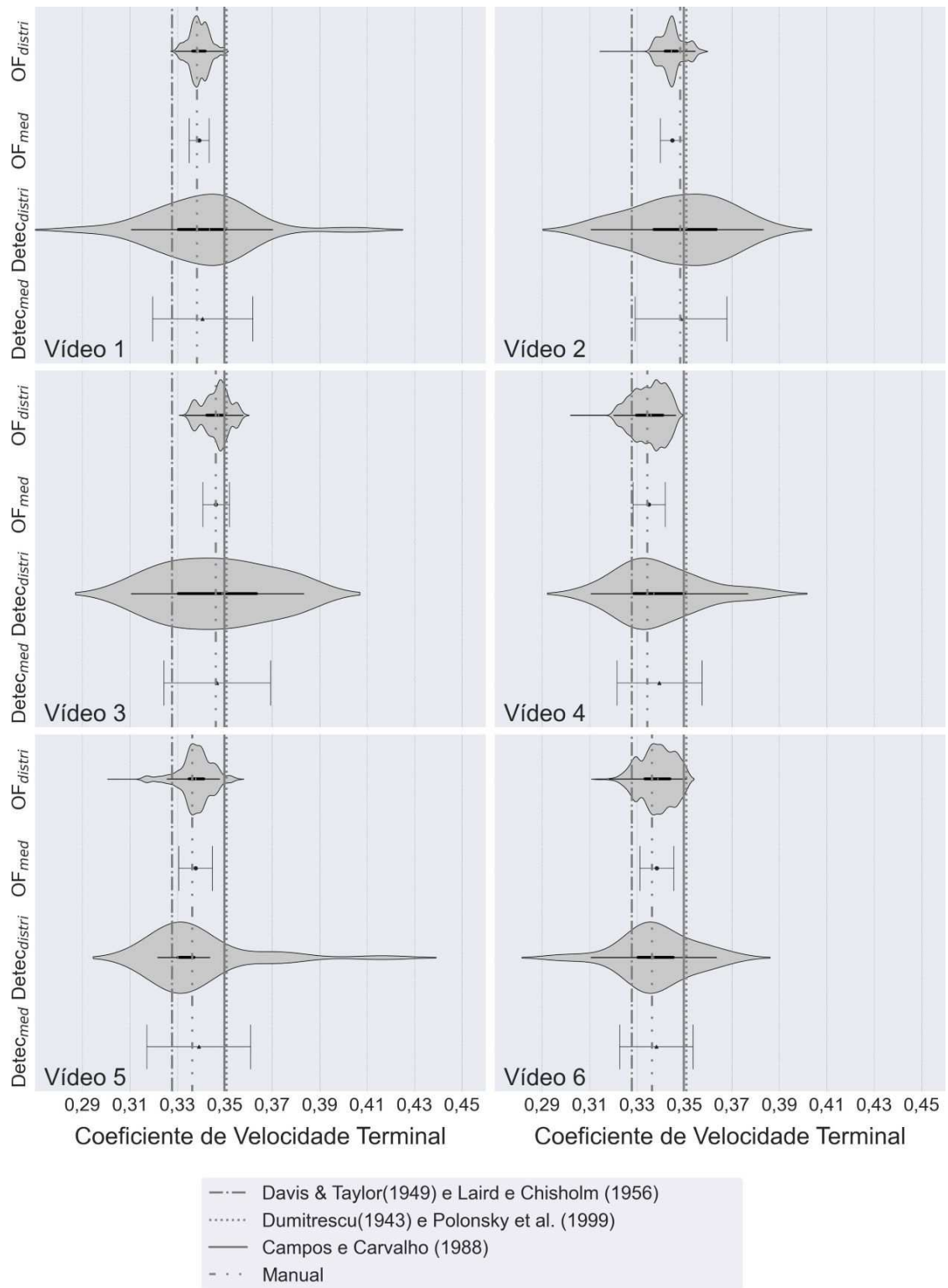
Ao avaliar a dispersão através do gráfico de violino para as velocidades, observa-se que os métodos de detecção ($\text{Detec}_{\text{distr}_i}$) apresentam uma dispersão significativamente maior, resultando em valores mais heterogêneos e introduzindo incertezas na comparação

Figura 24: Avaliação dos vetores de deslocamento pelos métodos de detecção e fluxo óptico utilizando dispersão em violino e média com um desvio padrão.



dos dados. Este fenômeno pode estar relacionado ao fato de que o detector não possui precisão exata na delimitação dos pixels que compõem o contorno da bolha. Durante a seleção das caixas de detecção, pode haver inclusão de pixels que não pertencem à bolha,

Figura 25: Avaliação de Coeficientes de Velocidade Terminal pelos métodos de detecção e fluxo óptico utilizando dispersão em violino e média com um desvio padrão.



o que difere de um método de treinamento por segmentação, no qual o contorno do objeto é delimitado pixel a pixel. Esta imprecisão na detecção pode explicar a maior dispersão observada nos dados de velocidade.

Tabela 5: Dado obtidos a partir do algoritmo pelo método de detecção.

Variáveis		Vídeos					
		1	2	3	4	5	6
$v_{BT,detec}$ (m/s)	Média	0,170	0,177	0,17	0,173	0,170	0,172
	σ	0,010	0,009	0,010	0,009	0,010	0,008
	1Q	0,168	0,171	0,168	0,167	0,168	0,168
	Mediana	0,175	0,178	0,177	0,171	0,168	0,171
	3Q	0,178	0,185	0,185	0,178	0,171	0,176
$C_{t,detec}$	Média	0,340	0,350	0,350	0,340	0,340	0,340
	σ	0,020	0,020	0,020	0,020	0,020	0,020
	1Q	0,331	0,337	0,331	0,329	0,331	0,331
	Mediana	0,344	0,350	0,349	0,337	0,331	0,337
	3Q	0,350	0,364	0,364	0,350	0,337	0,345
$Acurácia_{BT,detec}$		90%	91%	91%	91%	91%	91%

Assim, ao identificar novas bolhas, é provável que sua delimitação herde essas imprecisões do treinamento, não fornecendo uma representação precisa da extremidade da bolha e, conseqüentemente, introduzindo desvios no cálculo da velocidade.

Por outro lado, para o método que utiliza fluxo óptico (OF_{distri}), as dispersões são menores. No entanto, é importante notar a presença de alguns pontos mais distantes dos limites inferiores e superiores, como nos vídeos 1 e 5. Nestes casos, a estrutura de violino apresenta uma região de dispersão de baixa densidade de valores, sendo qualitativamente mais estreitas. Isso pode ter sido causado por aproximações do algoritmo que erroneamente identificaram dois pixels consecutivos que, na realidade, não estavam correlacionados. Tais problemas podem ser atribuídos à variação na iluminação da imagem e aos movimentos dos contornos da bolha, que podem não ser contínuos devido ao fluxo do fluido.

Essas análises podem ser mais esclarecedoras ao considerar as percentagens significativas do desvio padrão em relação à média e a distância entre o 1º e o 3º quartis em relação à mediana. Portanto, avaliar essas dispersões de maneira percentual em relação ao valor avaliado é fundamental.

Além disso, é crucial considerar o impacto dessas dispersões na interpretação dos resultados. Desvios padrão mais elevados indicam uma maior variabilidade nos dados, o que pode afetar a confiabilidade das conclusões obtidas a partir da análise. Portanto, ao interpretar os resultados, é fundamental ter em mente não apenas os valores médios, mas também a dispersão dos dados.

Avaliando a relação percentual entre o desvio padrão e as médias encontradas para

Tabela 6: Dado obtidos a partir do algoritmo pelo método de fluxo óptico e manual.

Variáveis		Vídeos					
		1	2	3	4	5	6
$v_{BT, \text{óptico}}(\text{m/s})$	Média	0,172	0,175	0,176	0,170	0,172	0,172
	σ	0,002	0,003	0,003	0,003	0,004	0,004
	1Q	0,171	0,174	0,174	0,168	0,170	0,169
	Mediana	0,172	0,175	0,177	0,171	0,172	0,172
	3Q	0,174	0,176	0,178	0,173	0,173	0,175
$C_{t, \text{óptico}}$	Média	0,339	0,345	0,346	0,335	0,338	0,338
	σ	0,004	0,005	0,006	0,007	0,007	0,007
	1Q	0,337	0,342	0,343	0,330	0,335	0,333
	Mediana	0,339	0,345	0,348	0,336	0,338	0,339
	3Q	0,342	0,347	0,350	0,341	0,341	0,344
$v_{BT, \text{manual}}(\text{m/s})$		0,172	0,177	0,176	0,170	0,171	0,171
$C_{t, \text{manual}}$		0,338	0,348	0,346	0,334	0,336	0,336

Tabela 7: Dados teóricos da velocidade terminal.

Referência	$C_{t, \text{teórico}}$	$V_{BT, \text{teórico}}(\text{m/s})$
Dumitrescu (1943) & Polonsky et al. (1999)	0,351	0,1783
Davies e Taylor (1950) & Laird e Chisholm (1956)	0,328	0,1666
Campos e Carvalho (1988)	0,350	0,1778

cada velocidade calculada, observa-se que o método de detecção apresentou um desvio padrão médio de 6,3% em relação aos dados do vídeo 5, enquanto o método óptico registrou 2,1% nos vídeos 5 e 6. Esses valores indicam que o método óptico proporciona uma variação menor em torno das médias de velocidade calculadas, sugerindo uma maior consistência nos resultados. Isso significa que as medições de velocidade obtidas pelo método óptico são mais próximas umas das outras, o que é crucial para uma análise precisa e confiável.

Ao avaliar a dispersão percentual entre os 1º e 3º quartis em relação à mediana, os maiores valores encontrados foram de 5,3% para o método de detecção (vídeo 3) e 1,8% para o método óptico (vídeo 3). Esses resultados destacam que o método óptico apresenta uma dispersão menor em torno da mediana das velocidades calculadas, indicando uma distribuição mais concentrada dos dados.

Essa menor dispersão nos resultados do método óptico é visualmente evidenciada pela Figura 22, onde é possível observar que os valores de velocidade calculados estão

mais próximos da linha central em comparação com o método de detecção. Isso reforça a conclusão de que o método óptico é mais eficaz em fornecer medições de velocidade mais consistentes e precisas, mesmo em condições variáveis como variações na iluminação e movimentos dos contornos das bolhas.

Este mesmo comportamento se estende para as constantes de velocidade terminal, uma vez que estas foram calculadas utilizando a velocidade juntamente com a constante gravitacional e o diâmetro da tubulação. Sendo estes dois últimos constantes e a constante diretamente proporcional à velocidade, os comportamentos dos gráficos apresentados nas Figuras 24 e 25 são semelhantes.

6.3.3 Comparação com a Literatura

Para a avaliação comparativa dos dados em relação à literatura, foram confrontadas as médias e medianas de cada um dos modelos com os valores teóricos apresentados na Tabela 8, evidenciando as diferenças percentuais entre os dados coletados. Essas diferenças percentuais foram calculadas pela razão entre o valor absoluto da diferença entre o valor teórico e o valor calculado e o valor teórico. Além disso, a mesma diferença percentual foi calculada comparando os métodos de detecção e fluxo óptico, utilizando como referência o método manual.

6.3.4 Comparação com a Literatura

A análise das velocidades mostra que as diferenças percentuais para ambos os casos não excederam 6%. Ao comparar os dados com a Figura 24, observa-se que as médias e medianas geralmente se encontram dentro da faixa dos valores teóricos, que variam entre 0,1666 m/s e 0,1783 m/s, os quais são calculados utilizando valores das constantes terminais amplamente aceitos na literatura.

A apresentação de uma faixa de valores teóricos demonstra que variações podem ocorrer devido a diferentes aproximações matemáticas ou condições experimentais variáveis, como temperatura, pressão e materiais da tubulação. Tais variações podem resultar em pequenas discrepâncias entre os valores, permitindo a apresentação de uma faixa de valores observáveis no experimento.

Embora os mesmos fluidos da literatura (água-ar) tenham sido utilizados e a luminosidade, posição da câmera e etapas experimentais tenham sido mantidas consistentes, os valores de velocidade não permaneceram constantes entre os vídeos. Essa inconsistência foi observada tanto nos resultados computacionais quanto no método manual de análise.

Além das variações mencionadas, outros fatores podem contribuir para a variação nos valores experimentais, como a presença de impurezas no fluido ou variações na concentração de partículas. No entanto, é importante ressaltar que tais variáveis provavelmente

Tabela 8: Comparação percentual entre os valores teóricos e manuais com as médias e medianas das velocidades obtidas pelos métodos de detecção e fluxo óptico.

Variáveis		Vídeos					
		1	2	3	4	5	6
Dumitrescu (1943) & Polonsky et al. (1999): 0,1783 m/s							
$v_{BT,detec}$	Média	2,9%	0,7%	1,2%	3,3%	3,4%	3,6%
	Mediana	2,9%	0,7%	1,2%	3,3%	3,4%	3,6%
$v_{BT,ótico}$	Média	3,3%	1,7%	1,3%	4,5%	3,8%	3,6%
	Mediana	3,4%	1,8%	0,9%	4,3%	3,7%	3,4%
Davies e Taylor (1950) & Laird e Chisholm (1956): 0,1666 m/s							
$v_{BT,detec}$	Média	4,0%	5,7%	3,4%	3,1%	3,1%	3,2%
	Mediana	2,8%	4,8%	1,8%	2,8%	0,8%	1,8%
$v_{BT,ótico}$	Média	2,6%	4,8%	2,4%	2,1%	1,5%	2,0%
	Mediana	2,8%	4,7%	2,7%	2,1%	1,8%	2,2%
Campos e Carvalho (1988): 0,1778 m/s							
$v_{BT,detec}$	Média	2,5%	0,9%	3,1%	3,3%	3,3%	3,3%
	Mediana	3,6%	1,8%	4,6%	3,6%	5,5%	4,6%
$v_{BT,ótico}$	Média	3,4%	5,2%	5,6%	2,2%	3,0%	3,2%
	Mediana	3,4%	5,1%	6,0%	2,5%	3,1%	3,4%
Manual (m/s):		0,172	0,177	0,176	0,170	0,171	0,171
$v_{BT,detec}$	Média	0,7%	0,1%	0,1%	1,5%	0,8%	0,5%
	Mediana	1,6%	0,6%	0,8%	0,8%	1,8%	0,2%
$v_{BT,ótico}$	Média	0,2%	1,0%	0,0%	0,2%	0,4%	0,6%
	Mediana	0,1%	1,0%	0,4%	0,4%	0,4%	0,7%

não alteram os valores de forma significativa no caso específico de bolhas de Taylor em meio infinito, pois o comportamento das bolhas é, neste contexto, predominantemente determinado pelas propriedades físicas do fluido e pela geometria do sistema experimental.

Um fator relevante a ser considerado é o possível efeito de distorção da câmera devido à curvatura da lente, que pode afetar a perspectiva da imagem final e, consequentemente, a proporção entre o diâmetro em pixels e o diâmetro real em metros. Essa variável é sensível no modelo, pois sua alteração afeta todos os valores coletados pelo experimento.

Apesar dessas considerações, os valores encontrados ainda se mantêm dentro de uma faixa de até 6% em relação aos valores teóricos, permitindo inferir que o modelo consegue estimar uma velocidade próxima do valor teórico. No entanto, devido à distribuição dos dados de velocidade, o método de detecção pela média ou mediana, por ser mais disperso,

pode ser mais uma estimativa da ordem de grandeza do que do valor exato. Isso ocorre porque, devido a imprecisões associadas à resolução de alguns pixels, pode haver uma variação entre os quadros que não reflete a real velocidade da bolha.

Por outro lado, o método de fluxo óptico, integrado ao detector, oferece uma representação mais precisa do movimento das bolhas ao longo do tempo, com especial ênfase na avaliação da ponta da bolha, onde se encontra a velocidade de interesse. Além disso, por não depender de um processo de treinamento e basear-se apenas na avaliação de parâmetros luminotécnicos, o método é menos suscetível a erros associados à interpretação experimental, o que pode afetar os dados de treinamento.

Para tentar comparar os dados obtidos pelo algoritmo com as velocidades reais das bolhas, utilizou-se o método manual como uma aproximação da realidade, com um experimentalista aferindo a ponta da bolha em três instantes do vídeo (início, meio e fim da passagem da bolha de Taylor). Esperava-se encontrar um valor mais próximo do real, mas vale ressaltar que essa medição também depende da proporção pixel/metros.

Avaliando a Tabela 8, comparando o método manual aos valores obtidos pelos métodos de detecção e fluxo óptico, as diferenças não ultrapassam 1,8%, sendo ainda menores quando considerados os valores obtidos pelo método de fluxo óptico, que não ultrapassam 1,0%. Esse fator demonstra que o método automático se alinha bem com uma medição manual utilizando filmagens, embora a comparação seja proporcional em termos de velocidade em pixels por segundo.

Dessa forma, é provável que a divergência dos valores medidos com os teóricos possa ser atribuída a uma ou mais das hipóteses citadas anteriormente, como pequenos erros na calibração da proporção pixel/metros, distorção da câmera, presença de impurezas e diferenças experimentais e teóricas entre os modelos e o experimento realizado.

Contudo, os valores ainda se apresentam dentro de uma faixa aceitável para serem considerados boas estimativas de uma medição real de uma bolha de Taylor, a qual pode apresentar flutuações e desviar do teórico esperado, promovendo uma medição automática mais próxima da realidade, mesmo considerando fatores que causem divergências em relação ao valor teórico.

6.3.5 Velocidade de Processamento do Algoritmo

Em todos os vídeos analisados foi medido o tempo de processamento e comparado com o tempo original do vídeo, obtendo os resultados da Tabela 9. Além disso também foram medidos o número de quadros em que objetos foram identificados, e o número total de vetores de deslocamento entre um quadro e outro. Vale ressaltar que os vetores foram calculados entre um quadro e outro.

Os resultados demonstram que, na maioria dos casos, o tempo de processamento do algoritmo foi maior do que o tempo de reprodução dos vídeos, exceto no vídeo 2

Tabela 9: Avaliação da Velocidade de Processamento do Algoritmo em Relação ao Tempo Real do Vídeo

Variáveis	Vídeos					
	1	2	3	4	5	6
Tempo do Vídeo (s)	5,88	7,23	4,68	5,53	5,18	4,88
Tempo do Algoritmo (s)	6,33	7,13	4,95	7,01	5,64	5,92
Diferença	+7,58%	-1,48%	+5,66%	+26,69%	+8,88%	+21,32%
Velocidade (FPS)	56	61	57	47	55	49
Nº de Quadros com Objeto	83	78	78	85	80	84
Nº de Pontos OF Detectados	1914	1466	1416	1966	1593	2094

onde o tempo do algoritmo foi ligeiramente menor (-1,48%). Esta variação no tempo de processamento depende diretamente do número de objetos detectados e da quantidade de pontos avaliados pelo método de fluxo óptico. À medida que aumenta o número de detecções, também aumenta o número de vetores de deslocamento calculados, o que eleva o esforço computacional e, conseqüentemente, o tempo de processamento.

Analisando a Tabela 9, observa-se que os vídeos com maior número de pontos detectados, como os vídeos 4 e 6, apresentaram uma diferença significativa no tempo de processamento (+26,69% e +21,32%, respectivamente). Isso reflete a relação direta entre o número de pontos de detecção e de número de quadro avaliados ao esforço computacional necessário.

Além disso, a velocidade de processamento, medida em frames por segundo (FPS), variou de 47 a 61 FPS. Embora o método de fluxo óptico tenha demonstrado eficiência em termos de precisão, a quantidade de detecções impacta diretamente no tempo de processamento, especialmente em vídeos com muitos objetos detectados.

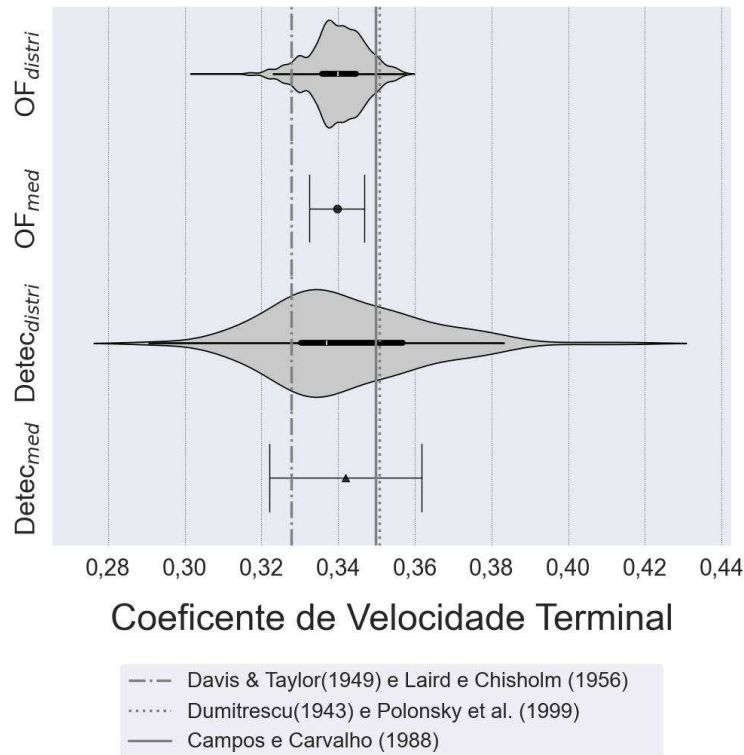
A comparação entre o tempo de processamento do algoritmo e o tempo do vídeo evidencia que, apesar de o processamento ser geralmente mais lento que a reprodução do vídeo, a metodologia aplicada permite a análise em tempo quase real. Esta capacidade é essencial para aplicações práticas onde a velocidade de processamento é crucial. No entanto, é evidente que a otimização do algoritmo poderia focar na redução do número de pontos avaliados, ou na implementação de técnicas de paralelização para melhorar o desempenho computacional.

Em resumo, os resultados confirmam a eficácia do método, embora destaquem a necessidade de balancear o número de detecções com o esforço computacional. A aplicação delimitada do método de fluxo óptico, como discutido previamente, se mostrou essencial para reduzir o esforço computacional, mantendo a precisão das medições. Este equilíbrio é fundamental para a aplicabilidade prática do algoritmo em cenários industriais, onde tanto a precisão quanto a velocidade de processamento são necessárias.

6.3.6 Constante de Velocidade Terminal

Com o objetivo de avaliar um coeficiente de velocidade terminal para o experimento como um todo, todos os dados de velocidade foram utilizados para calcular uma constante única. Para esse fim, foi realizada uma análise da distribuição dos dados para o detector e o método do fluxo óptico, conforme apresentado na Figura 26, juntamente com os valores obtidos pela média com desvio padrão e a avaliação dos quartis, conforme mostrado na Tabela 10.

Figura 26: Avaliação do coeficiente de velocidade terminal de todos os vídeos pelo método de detecção utilizando dispersão em violino e média com um desvio padrão.



Conforme esperado, a distribuição dos valores para o método do fluxo óptico permanece menor, refletindo o comportamento individual de cada vídeo e, portanto, será o método escolhido para mensurar a constante de velocidade terminal do experimento.

O valor obtido para a constante utilizando o método óptico, tanto pela média quanto pela mediana, foi de 0,340, posicionando-se como um valor intermediário entre os extremos

Tabela 10: Dados médios da constante de velocidade terminal obtidos a partir de todos os vídeos.

Variáveis		Valores Gerais
$C_{t, \text{detec}}$	Média	0,34
	σ	0,02
	1Q	0,330
	2Q	0,337
	3Q	0,357
$C_{t, \text{óptico}}$	Média	0,340
	σ	0,007
	1Q	0,336
	2Q	0,340
	3Q	0,345

apresentados na literatura. Essa medida, juntamente com a avaliação do desvio padrão, cobre uma parte significativa da faixa entre esses valores, demonstrando-se satisfatória do ponto de vista conceitual.

No entanto, é crucial ressaltar que a interpretação dos resultados deve considerar a complexidade do sistema experimental. A variabilidade nos dados pode ser influenciada por diversos fatores, incluindo variações nas condições ambientais, na preparação da amostra e nas técnicas de análise utilizadas.

Ademais, é fundamental levar em conta as limitações e incertezas associadas às técnicas de medição empregadas. A calibração precisa dos equipamentos, a precisão na identificação das bolhas e a validade das suposições subjacentes aos modelos utilizados podem impactar significativamente os resultados obtidos.

7 Conclusão

Este estudo investigou a aplicação de técnicas avançadas de visão computacional para a detecção, rastreamento e medição da velocidade de bolhas de Taylor em escoamentos verticais líquido-gás. A eficácia dos métodos foi validada por comparação com dados coletados a partir das filmagens experimentais e com valores da literatura, demonstrando a robustez e a confiabilidade do sistema desenvolvido.

O modelo YOLOv8 mostrou-se eficaz na detecção das bolhas ao longo dos quadros, permitindo acompanhar sua trajetória ao longo do tempo, com uma precisão média de 84,3% e uma acurácia média de 88%. A detecção também permitiu a extração do comprimento das bolhas, com medições que, na maioria dos casos, estavam próximas dos valores calculados manualmente. As flutuações ocorreram principalmente na região da base de bolhas muito longas, indicando uma limitação do modelo. Essa limitação foi percebida ao comparar os dados do detector com os métodos manuais; nos vídeos analisados, as médias e medianas estavam bastante próximas dos valores manuais, com desvios máximos de 1,1% e 1,9%, respectivamente. No entanto, no vídeo 2, observou-se desvios maiores (7,4% na média e 9,5% na mediana), que possuía o maior comprimento de bolha entre os demais, o que possivelmente se deve à dificuldade do detector em capturar bolhas grandes que ocupam uma grande parte da tubulação. Apesar dessas limitações, o detector demonstrou capacidade de rastrear as bolhas de Taylor ao longo do tempo, sugerindo que a utilização de um conjunto de dados mais diversificado poderia melhorar ainda mais o desempenho do modelo.

A integração do método de fluxo óptico com o YOLOv8 resultou em uma abordagem não intrusiva e eficaz para medir a velocidade das bolhas. A medição da velocidade foi realizada utilizando dois métodos distintos: rastreamento com o detector YOLOv8 e rastreamento utilizando fluxo óptico. Ambos os métodos produziram resultados consistentes com os encontrados na literatura e com os dados manuais, sendo que o método de fluxo óptico se destacou por sua menor dispersão e maior consistência dos resultados. As diferenças percentuais entre os dados obtidos e os valores da literatura não excederam 6%, situando-se dentro da faixa teórica de 0,167 m/s a 0,178 m/s. Comparando com o método manual, os dados do método de fluxo óptico apresentaram menor dispersão (diferenças inferiores a 1,0%), enquanto os dados do método de detecção apresentaram diferenças de até 1,8%. Esses resultados indicam que o método automático está bem alinhado com as medições manuais, corroborando sua funcionalidade e confiabilidade.

A análise estatística reforçou a superioridade do método de fluxo óptico em termos de consistência e menor dispersão dos dados. A relação percentual entre o desvio padrão e a média foi de 6,3% para o método de detecção e 2,1% para o método óptico. O desvio percentual entre a mediana e os quartis foi de 5,3% para o método de detecção e 1,8% para

o método de fluxo óptico. Esses resultados confirmam que o método de fluxo óptico é mais confiável e preciso para medir a velocidade das bolhas, quando comparado à utilização isolada do detector.

Assim, os métodos desenvolvidos mostraram-se viáveis e fornecem uma ferramenta aplicável para estudos em dinâmica de bolhas e aplicações correlatas. A metodologia aplicada apresentou grande potencial para o estudo de comportamentos de bolhas de Taylor e outros tipos de escoamento com estruturas bem definidas, possibilitando avanços significativos na compreensão e controle de escoamentos bifásicos. Esta abordagem oferece uma base para pesquisas futuras e para a otimização de processos envolvendo escoamentos translúcidos. No entanto, recomenda-se uma investigação mais detalhada sobre as limitações observadas, especialmente em cenários com bolhas maiores, a fim de aprimorar ainda mais a precisão e a confiabilidade dos métodos propostos.

7.1 Trabalhos Futuros

Embora o presente trabalho tenha atingido seus objetivos, várias áreas para desenvolvimento futuro podem ser exploradas, a fim de aprimorar ainda mais os resultados obtidos.

- Aprimoramento do banco de dados:** A ampliação do banco de dados utilizado para o treinamento do modelo é uma estratégia fundamental para aumentar a precisão da detecção. A inclusão de uma maior variedade de dimensões de bolhas de Taylor nos dados de treinamento pode melhorar a capacidade do modelo de lidar com diferentes condições experimentais. Atualmente, o sistema apresenta limitações na captura de bolhas que ocupam grandes áreas do campo de visão da câmera, especialmente aquelas cujo comprimento ultrapassa a extensão do campo de visão da filmagem. Isso ocorre porque bolhas de maior comprimento, ao ocupar quase toda a extensão do campo de visão, dificultam a detecção precisa de suas extremidades. Portanto, a inclusão de bolhas com diversos tamanhos pode aprimorar a capacidade de detecção, permitindo que o modelo seja mais adaptável a cenários experimentais variados. Além disso, a diversificação dos dados pode envolver a análise de bolhas em diferentes estágios de alongamento, o que pode influenciar a dinâmica de detecção e a precisão das medições de velocidade, principalmente para bolhas em estados próximos da formação de caudas longas, que possuem características particulares de movimento.
- Desenvolvimento de detector para bolhas maiores:** Uma limitação observada foi a dificuldade em medir bolhas de maior comprimento, que ocupam praticamente todo o campo de visão da câmera, dificultando a análise precisa de suas extremidades. Para superar esse desafio, uma possível solução seria a implementação de um detector capaz de identificar tanto a ponta quanto a cauda da bolha de Taylor. A partir dessa detecção, seria possível medir o deslocamento da bolha desde o momento em que a ponta começa a atravessar uma posição fixa na tubulação até o momento em que a cauda a ultrapassa. Esse processo permitiria a medição da velocidade da bolha ao longo do seu comprimento, utilizando diferentes pontos de observação. Uma abordagem possível seria fixar a posição de um ponto na tubulação e criar um detector que acompanhasse a ponta e a cauda da bolha durante seu deslocamento, obtendo informações sobre a velocidade com base na passagem das extremidades da bolha por esse ponto fixo. Essa metodologia poderia ser estendida para calcular a velocidade média da bolha utilizando múltiplos pontos de passagem, fornecendo uma medição mais detalhada do movimento da bolha ao longo da tubulação.
- Extensão para escoamentos pistonados:** A aplicação dos métodos para a

análise de bolhas em escoamentos do tipo pistonado, onde múltiplas bolhas se movem de forma contínua e interagem entre si, representa um possível avanço. Nesse caso, seria possível estudar a dinâmica de bolhas consecutivas e sua interação com o fluido circundante. A detecção e o rastreamento das bolhas permitiriam não apenas medir sua velocidade, mas também observar como as bolhas afetam o perfil do fluido e as interações entre as bolhas durante o movimento. A análise das esteiras de bolhas, utilizando o método de fluxo óptico, poderia proporcionar informações sobre os vetores de velocidade na região da esteira das bolhas, permitindo uma melhor compreensão das dinâmicas de escoamento em diferentes regimes de fluxo. Esses dados poderiam fornecer uma visão mais detalhada sobre o impacto das bolhas no fluido e auxiliar no desenvolvimento de modelos para escoamentos mais complexos.

- **Estudos em diferentes condições de fluxo:** Ampliar os estudos para incluir diferentes condições de fluxo e tipos de fluido pode fornecer informações valiosas sobre a adaptação das bolhas a diferentes ambientes experimentais. Variações em parâmetros como viscosidade, densidade, taxa de fluxo e outros fatores que alteram as características do fluido podem modificar significativamente a dinâmica das bolhas. Estudar essas variações possibilitará o desenvolvimento de modelos mais generalizáveis, que poderão ser aplicados a uma gama mais ampla de situações, incluindo aquelas que envolvem a interação das bolhas com obstáculos e outros fenômenos presentes em sistemas industriais. Tais estudos podem contribuir para a adaptação de tecnologias que envolvem escoamentos bifásicos em contextos industriais, como processos de separação de fases ou controle de reações em leitos fluidizados.

A implementação dessas melhorias, juntamente com o desenvolvimento de novas abordagens para a análise das bolhas, proporcionará uma compreensão mais profunda das dinâmicas de escoamentos bifásicos, possibilitando avanços significativos na aplicação das técnicas de visão computacional e medição de velocidade em sistemas complexos de fluido.

Referências

- ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. In: IEEE. *2017 international conference on engineering and technology (ICET)*. 2017. p. 1–6. Disponível em: <https://doi.org/10.1109/ICEngTechnol.2017.8308186>.
- AMERAN, H. et al. Multiphase flow velocity measurement of chemical processes using electrical tomography: A review. *2016 IEEE Conference on Systems, Process and Control (ICSPC)*, n. 24, p. 130–135, 2016. Disponível em: <https://doi.org/10.1109/SPC.2016.7920717>.
- ANSARI, S. *Building Computer Vision Applications Using Artificial Neural Networks*. Apress, 2020. Disponível em: <https://doi.org/10.1007/978-1-4842-9866-4>.
- ARIF, R. B. et al. Study and observation of the variations of accuracies for handwritten digits recognition with various hidden layers and epochs using convolutional neural network. In: IEEE. *2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT)*. 2018. p. 112–117. Disponível em: <https://doi.org/10.48550/arXiv.1809.06187>.
- BEAUCHEMIN, S. S.; BARRON, J. L. The computation of optical flow. *ACM computing surveys (CSUR)*, v. 27, n. 3, p. 433–466, 1995. Disponível em: <https://doi.org/10.1145/212094.212141>.
- BLAND, M. Estimating mean and standard deviation from the sample size, three quartiles, minimum, and maximum. *International Journal of Statistics in Medical Research, Lifescience Global*, v. 4, n. 1, p. 57, 2015. Disponível em: <http://dx.doi.org/10.6000/1929-6029.2015.04.01.6>.
- BOCHKOVSKIY, A.; WANG, C.-Y.; LIAO, H.-Y. M. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020. Disponível em: <https://doi.org/10.48550/arXiv.2004.10934>.
- CAMPOS, J.; CARVALHO, J. G. D. An experimental study of the wake of gas slugs rising in liquids. *Journal of Fluid Mechanics*, Cambridge University Press, v. 196, p. 27–37, 1988. Disponível em: <https://doi.org/10.1017/S0022112088002599>.
- CHEN, W.-C. et al. A framework for real-time vehicle counting and velocity estimation using deep learning. *Sustainable Computing: Informatics and Systems*, Elsevier, v. 40, p. 100927, 2023. Disponível em: <https://doi.org/10.1016/j.suscom.2023.100927>.
- DANELLI, M. et al. Eco: Efficient convolution operators for tracking. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [s.n.], 2017. p. 6638–6646. Disponível em: <https://doi.org/10.48550/arXiv.1611.09224>.
- DAVIES, R.; TAYLOR, G. I. 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*, The Royal Society London, v. 200, n. 1062, p. 375–390, 1950. Disponível em: <https://doi.org/10.1098/rspa.1950.0023>.

DEVASAGAYAM, J.; BOSMA, R.; COLLIER, C. M. A velocity program using the kanade-lucas-tomasi feature-tracking algorithm with demonstration for pressure and electroosmosis conditions. *Electrophoresis*, Wiley Online Library, v. 43, n. 7-8, p. 865–878, 2022. Disponível em: <https://doi.org/10.1002/elps.202100177>.

DING, T. et al. Transient flow vector distribution of cavitation bubbles in a vessel induced by pulsed focused ultrasound thrombolysis based on pyramid lk optical flow method. In: IEEE. *2021 IEEE International Ultrasonics Symposium (IUS)*. 2021. p. 1–4. Disponível em: <https://doi.org/10.1109/IUS52206.2021.9593868>.

DUMITRESCU, D. T. Strömung an einer luftblase im senkrechten rohr. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, Wiley Online Library, v. 23, n. 3, p. 139–149, 1943. Disponível em: <https://doi.org/10.1002/ZAMM.19430230303>.

FALCONE, G.; HEWITT, G. F.; ALIMONTI, C. *Multiphase Flow Meterring: Principles and applications*. [s.n.], 2009. v. 54. (Developments in Petroleum Science, v. 54). Disponível em: [https://doi.org/10.1016/S0376-7361\(09\)05403-X](https://doi.org/10.1016/S0376-7361(09)05403-X).

FERNANDES, R.; SEMIAT, R.; DUKLER, A. E. Hydrodynamic model for gas-liquid slug flow in vertical tubes. *AIChE journal*, Wiley Online Library, v. 29, n. 6, p. 981–989, 1983. Disponível em: <https://doi.org/10.1002/aic.690290617>.

GIRSHICK, R. Fast r-cnn. In: *Proceedings of the IEEE international conference on computer vision*. [s.n.], 2015. p. 1440–1448. Disponível em: <https://doi.org/10.48550/arXiv.1504.08083>.

HARRIS, C.; STEPHENS, M. et al. A combined corner and edge detector. In: CITESEER. *Alvey vision conference*. 1988. v. 15, n. 50, p. 10–5244. Disponível em: <https://doi.org/10.5244/C.2.23>.

HINTZE, J. L.; NELSON, R. D. Violin plots: a box plot-density trace synergism. *The American Statistician*, Taylor & Francis, v. 52, n. 2, p. 181–184, 1998.

HORN, B. K.; SCHUNCK, B. G. Determining optical flow. *Artificial intelligence*, Elsevier, v. 17, n. 1-3, p. 185–203, 1981. Disponível em: [https://doi.org/10.1016/0004-3702\(81\)90024-2](https://doi.org/10.1016/0004-3702(81)90024-2).

HOUT, R. V. et al. Experimental investigation of the velocity field induced by a taylor bubble rising in stagnant water. *International Journal of Multiphase Flow*, Elsevier, v. 28, n. 4, p. 579–596, 2002. Disponível em: [https://doi.org/10.1016/S0301-9322\(01\)00082-9](https://doi.org/10.1016/S0301-9322(01)00082-9).

JOCHER, G.; CHAURASIA, A.; QIU, J. *YOLO by Ultralytics*. 2023. Disponível em: <https://github.com/ultralytics/ultralytics>.

JOCHER, G. et al. Yolov5. *Accessed March*, v. 7, p. 2021, 2020. Disponível em: <https://github.com/ultralytics/yolov5>.

KADISH, S. et al. Computer vision-based classification of flow regime and vapor quality in vertical two-phase flow. *Sensors*, MDPI, v. 22, n. 3, p. 996, 2022. Disponível em: <https://doi.org/10.3390/s22030996>.

KALE, K.; PAWAR, S.; DHULEKAR, P. Moving object tracking using optical flow and motion vector estimation. *2015 4th international conference on reliability, infocom technologies and optimization (ICRITO)(trends and future directions)*, p. 1–6, 2015. Disponível em: <https://doi.org/10.1109/ICRITO.2015.7359323>.

KANG, C.-W.; QUAN, S.; LOU, J. Numerical study of a taylor bubble rising in stagnant liquids. *Physical Review E*, APS, v. 81, n. 6, p. 066308, 2010. Disponível em: <https://doi.org/10.1103/PhysRevE.81.066308>.

KAPUTA, D. S.; LANDY, B. P. Yolbo: you only look back once—a low latency object tracker based on yolo and optical flow. *IEEE Access*, IEEE, v. 9, p. 82497–82507, 2021. Disponível em: <https://doi.org/10.1109/ACCESS.2021.3080136>.

KIM, K.; KIM, K.; JEONG, S. Application of yolo v5 and v8 for recognition of safety risk factors at construction sites. *Sustainability*, MDPI, v. 15, n. 20, p. 15179, 2023. Disponível em: <https://doi.org/10.3390/su152015179>.

LAIRD, A.; CHISHOLM, D. Pressure and forces along cylindrical bubbles in a vertical tube. *Industrial & Engineering Chemistry*, ACS Publications, v. 48, n. 8, p. 1361–1364, 1956. Disponível em: <https://doi.org/10.1021/ie50560a037>.

LI, C. et al. YOLOv6: A single-stage object detection framework for industrial applications. *arXiv preprint arXiv:2209.02976*, 2022. Disponível em: <https://doi.org/10.48550/arXiv.2209.02976>.

LI, L. et al. Leveraging the invariant side of dynamic trichomonas vaginalis via the fusion of optical flow. In: IEEE. *2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD)*. 2023. p. 668–672. Disponível em: <https://doi.org/10.1109/ICAIBD57115.2023.10206172>.

LIN, T.-Y. et al. Microsoft coco: Common objects in context. In: SPRINGER. *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13*. 2014. p. 740–755. Disponível em: <https://doi.org/10.48550/arXiv.1405.0312>.

LIU, K. et al. A real-time method to estimate speed of object based on object detection and optical flow calculation. In: IOP PUBLISHING. *Journal of Physics: Conference Series*. 2018. v. 1004, n. 1, p. 012003. Disponível em: <https://doi.org/10.1088/1742-6596/1004/1/012003>.

LIU, W. et al. Ssd: Single shot multibox detector. In: SPRINGER. *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. 2016. p. 21–37. Disponível em: <https://doi.org/10.48550/arXiv.1512.02325>.

LIVINGSTON, E. H. The mean and standard deviation: what does it all mean? *Journal of Surgical Research*, Elsevier, v. 119, n. 2, p. 117–123, 2004. Disponível em: <https://doi.org/10.1016/j.jss.2004.02.008>.

LU, X. et al. Mimicdet: Bridging the gap between one-stage and two-stage object detection. In: SPRINGER. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIV* 16. 2020. p. 541–557. Disponível em: <https://doi.org/10.48550/arXiv.2009.11528>.

LUCAS, B. D.; KANADE, T. An iterative image registration technique with an application to stereo vision. In: *IJCAI'81: 7th international joint conference on Artificial intelligence*. [s.n.], 1981. v. 2, p. 674–679. Disponível em: <https://hal.science/hal-03697340>.

MAYOR, T. et al. Hydrodynamics of gas–liquid slug flow along vertical pipes in turbulent regime—an experimental study. *International Journal of Heat and Fluid Flow*, Elsevier, v. 29, n. 4, p. 1039–1053, 2008. Disponível em: <https://doi.org/10.1016/j.ijheatfluidflow.2008.02.013>.

MISRA, D. Mish: A self regularized non-monotonic activation function. *arXiv preprint arXiv:1908.08681*, 2019. Disponível em: <https://doi.org/10.48550/arXiv.1908.08681>.

NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*. [s.n.], 2010. p. 807–814. Disponível em: <https://www.cs.toronto.edu/~hinton/absps/reluICML.pdf>.

POLONSKY, S.; SHEMER, L.; BARNEA, D. The relation between the taylor bubble motion and the velocity field ahead of it. *International Journal of Multiphase Flow*, Elsevier, v. 25, n. 6-7, p. 957–975, 1999. Disponível em: [https://doi.org/10.1016/S0301-9322\(99\)00037-3](https://doi.org/10.1016/S0301-9322(99)00037-3).

QIMIN, X. et al. A methodology of vehicle speed estimation based on optical flow. In: IEEE. *Proceedings of 2014 IEEE International Conference on Service Operations and Logistics, and Informatics*. 2014. p. 33–37. Disponível em: <https://doi.org/10.1109/SOLI.2014.6960689>.

RAJAN, V.; RIDLEY, R.; RAFA, K. Multiphase flow measurement techniques - a review. *Journal of Energy Resources Technology*, p. 115–161, 1993. Disponível em: <https://doi.org/10.1115/1.2905987>.

RANE, N. Yolo and faster r-cnn object detection for smart industry 4.0 and industry 5.0: applications, challenges, and opportunities. *Available at SSRN 4624206*, 2023. Disponível em: <https://dx.doi.org/10.2139/ssrn.4624206>.

REDMON, J. et al. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [s.n.], 2016. p. 779–788. Disponível em: <https://doi.org/10.48550/arXiv.1506.02640>.

REDMON, J.; FARHADI, A. Yolo9000: better, faster, stronger. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [s.n.], 2017. p. 7263–7271. Disponível em: <https://doi.org/10.48550/arXiv.1612.08242>.

REDMON, J.; FARHADI, A. Yolo3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018. Disponível em: <https://doi.org/10.48550/arXiv.1804.02767>.

REN, S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, v. 28, 2015. Disponível em: <https://doi.org/10.48550/arXiv.1506.01497>.

SAKIB, S. et al. An overview of convolutional neural network: Its architecture and applications. Preprints, 2019. Disponível em: <http://dx.doi.org/10.20944/preprints201811.0546.v1>.

SCHÜTZE, H.; MANNING, C. D.; RAGHAVAN, P. *Introduction to information retrieval*. Cambridge University Press Cambridge, 2008. v. 39. Disponível em: <https://doi.org/10.1017/CBO9780511809071>.

SEONG, J. H. et al. Automated bubble analysis of high-speed subcooled flow boiling images using u-net transfer learning and global optical flow. *International Journal of Multiphase Flow*, Elsevier, v. 159, p. 104336, 2023. Disponível em: <https://doi.org/10.1016/j.ijmultiphaseflow.2022.104336>.

SHI, J. et al. Good features to track. In: IEEE. *1994 Proceedings of IEEE conference on computer vision and pattern recognition*. 1994. p. 593–600. Disponível em: <https://doi.org/10.1109/CVPR.1994.323794>.

SHOHAM, O. *Mechanistic modeling of gas-liquid two-phase flow in pipes*. 1. ed. Richardson, TX, USA: Society of Petroleum Engineers: [s.n.], 2006. Disponível em: <https://doi.org/10.2118/9781555631079>.

SU, L. et al. Transformer vibration detection based on yolov4 and optical flow in background of high proportion of renewable energy access. *Frontiers in Energy Research*, Frontiers, v. 10, p. 764903, 2022. Disponível em: <https://doi.org/10.3389/fenrg.2022.764903>.

TAITEL, Y.; BARNEA, D.; DUKLER, A. Modelling flow pattern transitions for steady upward gas-liquid flow in vertical tubes. *AIChE journal*, Wiley Online Library, v. 26, n. 3, p. 345–354, 1980. Disponível em: <https://doi.org/10.1002/aic.690260304>.

TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. *International conference on machine learning*. 2019. p. 6105–6114. Disponível em: <https://doi.org/10.48550/arXiv.1905.11946>.

TERVEN, J.; CÓRDOVA-ESPARZA, D.-M.; ROMERO-GONZÁLEZ, J.-A. A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction*, v. 5, n. 4, p. 1680–1716, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2304.00501>.

TIAN, Z. et al. Robust facial marker tracking based on a synthetic analysis of optical flows and the yolo network. *The Visual Computer*, Springer, p. 1–19, 2023. Disponível em: <https://doi.org/10.1007/s00371-023-02931-w>.

TORISAKI, S.; MIWA, S. Robust bubble feature extraction in gas-liquid two-phase flow using object detection technique. *Journal of Nuclear Science and Technology*, Taylor & Francis, v. 57, n. 11, p. 1231–1244, 2020. Disponível em: <https://doi.org/10.1080/00223131.2020.1779145>.

- VU, Q. D.; CHUNG, S.-T. Real-time robust human tracking based on lucas-kanade optical flow and deep detection for embedded surveillance. In: IEEE. *2017 8th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES)*. 2017. p. 1–6. Disponible em: <https://doi.org/10.1109/ICTEmSys.2017.7958761>.
- WANG, C.-Y.; BOCHKOVSKIY, A.; LIAO, H.-Y. M. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [s.n.], 2023. p. 7464–7475. Disponible em: <https://doi.org/10.48550/arXiv.2207.02696>.
- WANG, C.-Y.; LIAO, H.-Y. M.; YEH, I.-H. Designing network design strategies through gradient path analysis. *arXiv preprint arXiv:2211.04800*, 2022. Disponible em: <https://doi.org/10.48550/arXiv.2211.04800>.
- WANG, J. et al. Yolo-xray: A bubble defect detection algorithm for chip x-ray images based on improved yolov5. *Electronics*, MDPI, v. 12, n. 14, p. 3060, 2023. Disponible em: <https://doi.org/10.3390/electronics12143060>.
- XIAOWEI, W. Algorithm research on velocity measurement of biphasic flow. *Ninth International Conference on Hybrid Intelligent Systems*, n. 1, p. 471–474, 2009. Disponible em: <https://doi.org/10.1109/HIS.2009.211>.
- ZHANG, Z.; PENG, H. Deeper and wider siamese networks for real-time visual tracking. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [s.n.], 2019. p. 4591–4600. Disponible em: <https://doi.org/10.48550/arXiv.1901.01660>.
- ZHAO, Q.; ZHENG, C.; MA, W. An improved crucible spatial bubble detection based on yolov5 fusion target tracking. *Sensors*, MDPI, v. 22, n. 17, p. 6356, 2022. Disponible em: <https://doi.org/10.3390/s22176356>.
- ZHOU, W. et al. Bubble feature extraction in subcooled flow boiling using ai-based object detection and tracking techniques. *International Journal of Heat and Mass Transfer*, Elsevier, v. 222, p. 125188, 2024. Disponible em: <https://doi.org/10.1016/j.ijheatmasstransfer.2024.125188>.
- ZOU, Z. et al. Object detection in 20 years: A survey. *Proceedings of the IEEE*, IEEE, 2023. Disponible em: <https://doi.org/10.48550/arXiv.1905.05055>.