

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
ENGENHARIA DE CONTROLE E AUTOMAÇÃO

Ezequiel Eric Olejaz Freire - 219840

**Calibração de sistema de visão com
câmera de linha**

Porto Alegre
2023

Ezequiel Eric Olejaz Freire - 219840

Calibração de sistema de visão com câmera de linha

Trabalho de Conclusão de Curso (TCC-CCA) apresentado à COMGRAD-CCA da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título *Bacharel em Eng. de Controle e Automação*.

ORIENTADOR:

Prof. Dr. Pedro Rafael Bolognese Fernandes

CO-ORIENTADOR:

Dr. Carlos Solon Soares Guimarães Jr.

Porto Alegre
2023

Ezequiel Eric Olejaz Freire - 219840

Calibração de sistema de visão com câmera de linha

Trabalho de Conclusão de Curso (TCC-CCA) apresentado à COMGRAD-CCA da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título *Bacharel em Eng. de Controle e Automação*.

Orientador: _____
Prof. Dr. Pedro Rafael Bolognese Fernandes, UFRGS
Doutor pela Universidade Federal do Rio Grande do Sul
- Porto Alegre, Brasil

Co-orientador: _____
Dr. Carlos Solon Soares Guimarães Jr., ISI SIM
Doutor pela Universidade Federal do Rio Grande do Sul
- Porto Alegre, Brasil

Banca Examinadora:

Prof. Dr. Heraldo José de Amorim, UFRGS
Doutor pela Universidade Federal do Rio Grande do Sul - Porto Alegre, Brasil

Prof^ª. Dr^ª. Luciola Campestrini, UFRGS
Doutora pela Universidade Federal do Rio Grande do Sul - Porto Alegre, Brasil

Prof. Dr. Pedro Rafael Bolognese Fernandes, UFRGS
Doutor pela Universidade Universidade de Dortmund - Dortmund, Alemanha

Prof. Me. Alceu Heinke Frigeri
Coordenador de Curso
Eng. de Controle e Automação

Porto Alegre, setembro de 2023.

DEDICATÓRIA

Dedico esta obra ao Guilherme, pelos ensinamentos de gentileza para comigo mesmo sem os quais jamais a teria concluído.

AGRADECIMENTOS

À Universidade Federal do Rio Grande do Sul, pela formação técnica, cidadã e assistência estudantil.

Ao Instituto SENAI de Inovação - Sistemas Integrados em Metalmecânica, pelo ambiente propício à inovação, equipamentos e serviços disponibilizados e pelo incentivo ao desenvolvimento técnico-acadêmico.

Aos orientadores, pela disponibilidade e paciência constantes.

À família, pelo apoio incondicional e confiança inabalável.

Aos amigos, pela companhia na longa, árdua e transformadora jornada da graduação.

Aos presentes, que de perto ou à distância me ofereceram cuidado e compreensão com generosidade neste período que me foi tão caro entender o que estava acontecendo.

RESUMO

O trabalho prevê o desenvolvimento de um algoritmo de calibração de um sistema de visão composto por uma câmera de linha sobre uma esteira, incluindo parâmetros não-lineares de distorção da lente. A captura de imagens bidimensionais ocorre, nesse sistema, pelo movimento de objeto sobre a esteira, passando pelo campo de visão da câmera. Cada fotografia registra uma linha que, unidas de forma sequencial, formam uma representação 2D do objeto. O procedimento de calibração explora a identificação do modelo do sistema por Aprendizado de Máquina, permitindo descrever matematicamente a projeção de objetos 3D em uma imagem. A proposição é: parametrizar e montar o sistema de visão; projetar e confeccionar o padrão de calibração; implementar algoritmo de calibração com tecnologia de Aprendizado Profundo; realizar ensaios de calibração; avaliar os resultados em termos de convergência e erro de reprojeção; e revisão do método, discutindo seu impacto e enunciando possibilidades de desenvolvimento futuro.

Palavras-chave: Visão Computacional, Calibração, Sistemas de Visão, Câmera de linha, Aprendizado de Máquina, Redes Neurais, Aprendizado Profundo, Perceptron Multi-Camada.

ABSTRACT

The work envisages the development of a calibration algorithm for a vision system comprised of a line scan camera placed on a conveyor belt, encompassing non-linear lens distortion parameters. In this system, two-dimensional image capture occurs as objects move across the conveyor belt, traversing the camera's field of view. Each photograph records a line which, when sequentially joined, forms a 2D representation of the object. The calibration procedure leverages Machine Learning to identify the system's model, enabling the mathematical description of the projection of 3D objects onto an image. The proposition is to parameterize and assemble the vision system; design and create the calibration pattern; implement the calibration algorithm using Deep Learning technology; conduct calibration tests; evaluate the results in terms of convergence and reprojection error; and review the method, discussing its impact and outlining possibilities for future development.

Keywords: Machine Vision, Calibration, Vision System, Line-scan camera, Machine Learning, Neural Network, Deep Learning, Multi-Layer Perceptron.

Sumário

LISTA DE ILUSTRAÇÕES	10
LISTA DE TABELAS	11
LISTA DE ABREVIATURAS	11
1 INTRODUÇÃO	13
1.1 Objetivos	15
1.2 Estrutura	15
2 REVISÃO BIBLIOGRÁFICA E ESTADO DA ARTE	16
2.1 Câmeras de linha	17
2.2 Calibração de sistemas de visão	18
2.2.1 Métodos de calibração	19
2.2.2 Padrões de calibração	19
2.3 Modelo da câmera de linha	20
2.3.1 Distorção da lente	21
2.4 Aprendizado de máquina	22
2.4.1 Redes neurais	23
3 MATERIAIS E MÉTODOS	25
3.1 Padrão de referência	26
3.2 Ensaio de calibração	28
3.3 Sistema de visão	28
3.3.1 Preparo dos dados	29
3.3.2 Algoritmo de IA	29
4 RESULTADOS E DISCUSSÕES	31
4.1 Ensaio estático	31
4.2 Reconhecimento do padrão de calibração	32
4.3 Regressor MLP	33
5 CONSIDERAÇÕES FINAIS	36
REFERÊNCIAS	39

LISTA DE ILUSTRAÇÕES

1	Exemplo de aplicação industrial de sistema de visão com câmera de linha.	14
2	Exemplos de padrão de calibração de sistemas de visão.	19
3	Modelo de câmera estenopeico.	20
4	Representação gráfica da distorção radial.	22
5	Modelo do neurônio artificial Perceptron (esquerda) e uma rede neural profunda com três camadas totalmente conectadas (direita).	24
6	Padrão de calibração sobre base de deslocamento linear.	26
7	Sistema real de captura.	27
8	Imagem obtida no ensaio estático.	27
9	Imagem do padrão de calibração composto pelos tabuleiros 1 a 5, da esquerda para a direita.	28
10	Segmentação dos tabuleiros 1 a 5 em imagens separadas.	29
11	Registro do procedimento de calibração do padrão desenvolvido.	30
12	Detalhe da região central de imagem capturado no ensaio estático.	31
13	Curvas de cor, em escala de cinza, por linha de imagem do ensaio estático.	32
14	Curva de função de custo durante o treinamento.	33
15	Histogramas de erro absoluto por eixo.	34

LISTA DE TABELAS

4.1	Extração de pontos bem-sucedidas por imagem	32
4.2	Resultados estatísticos dos erros de reprojeção do conjunto de validação .	34

LISTA DE ABREVIATURAS

1D	<i>One-dimensional</i>	(Unidimensional)
2D	<i>Two-dimensional</i>	(Bidimensional)
3D	<i>Three-dimensional</i>	(Tridimensional)
CNN	<i>Convolutional Neural Network</i>	(Redes Neurais Convolucionais)
GHI	<i>Gated Recurrent Unit</i>	
LED	<i>Light Emitting Diode</i>	(Diodo Emissor de Luz)
LSTM	<i>Long Short-Term Memory</i>	(Memória Longa de Curto Prazo)
MSE	<i>Mean of squared error</i>	(Média do erro quadrático)
ML	<i>Machine Learning</i>	(Aprendizado de Máquina)
MLP	<i>Multi-Layer Perceptron</i>	(Perceptron Multi-Camadas)
RNN	<i>Recurrent Neural Network</i>	(Rede Neural Recorrente)
SVM	<i>Support Vector Machine</i>	(Máquinas de Vetores de Suporte)
VDC	<i>Volts Direct Current</i>	(Volts de Corrente Contínua)
VIM	Vocabulário Internacional de Metrologia	

1 INTRODUÇÃO

As câmeras de linha, do inglês *line scan cameras*, são elementos amplamente utilizados no âmbito da visão computacional (WANG et al., 2016). Com um sensor composto por uma única linha de pixels, essas câmeras têm a capacidade de capturar imagens com alta resolução e velocidade, demonstrando eficácia em ambientes industriais onde a análise e a inspeção de imagens são fatores determinantes.

Em aplicações industriais, as câmeras de linha desempenham papéis fundamentais. Na inspeção de produtos, elas garantem a qualidade na produção ao detectar defeitos, avaliar dimensões e verificar cores, proporcionando consistência e confiabilidade aos produtos. O monitoramento de processos é otimizado por meio do acompanhamento em tempo real dos procedimentos de fabricação, possibilitando ajustes e melhorias na eficiência.

Na classificação e seleção, essas câmeras são empregadas para distinguir objetos com base em características visuais, proporcionando rapidez e precisão na triagem. Além disso, a metrologia é beneficiada com a capacidade de medir dimensões e formas de objetos em movimento, essencial nas indústrias de cabos e fibras ópticas, por exemplo.

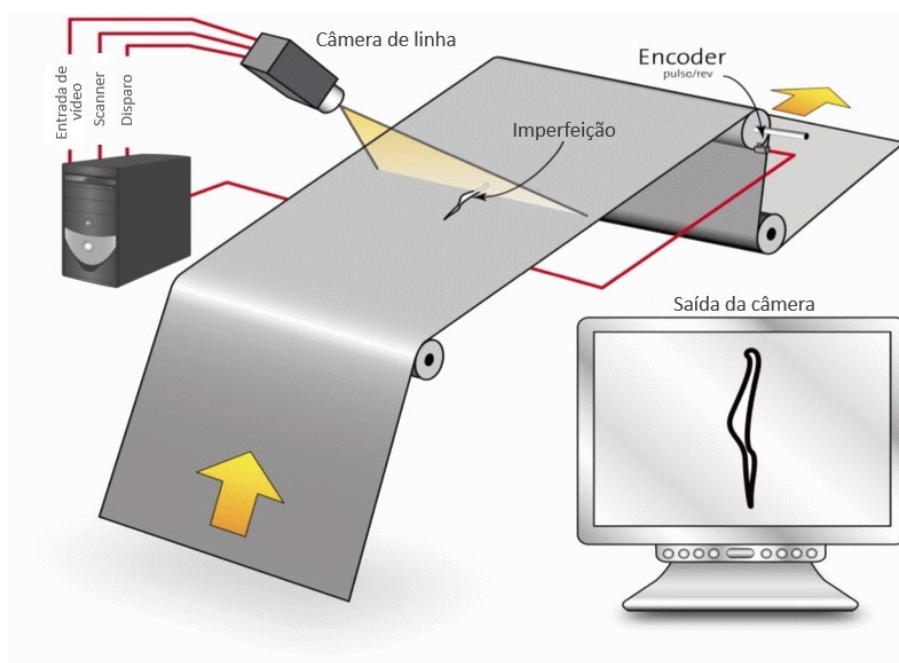
As características gerais de funcionamento da câmera de linha são (GONZALEZ et al., 2010):

1. Sensor de linha: As câmeras de linha incorporam sensores compostos por uma única linha de pixels. Essa configuração permite a captura de imagens de objetos em movimento contínuo.
2. Alta velocidade: Uma das vantagens dessas câmeras é a sua capacidade de adquirir imagens em elevadas taxas, muitas vezes alcançando milhares de linhas por segundo. Isso as torna ideais para cenários industriais com demandas de alta velocidade.
3. Alta resolução: A resolução da imagem é influenciada tanto pela quantidade de pixels no sensor quanto pela velocidade do objeto em movimento. Isso resulta em imagens que mantêm a qualidade mesmo em altas resoluções.
4. Iluminação uniforme: As câmeras de linha requerem uma fonte de luz uniforme e brilhante, geralmente fornecida por luzes do tipo diodo emissor de luz (do inglês *light emitting diode*, LED), para assegurar a qualidade das imagens capturadas.

5. Sincronização: A sincronização entre a câmera de linha e a velocidade do objeto em movimento é fundamental. Esse processo é frequentemente realizado usando *encoders* lineares ou rotativos.

Ao unir as capacidades das câmeras de linha com algoritmos avançados de visão computacional, é possível criar sistemas de inspeção automática de alta precisão e eficiência. Esses sistemas são versáteis, encontrando aplicação em diversos setores industriais, como manufatura, embalagem, alimentos e bebidas, eletrônicos, papel e celulose, entre outros. A Figura 1 ilustra uma operação de inspeção, na qual a câmera registra objetos se deslocando linearmente, de modo a gerar dados para diversas aplicações de visão computacional. Por exemplo, identificação de defeitos e características dimensionais.

Figura 1: Exemplo de aplicação industrial de sistema de visão com câmera de linha.



Stemmer Imaging, 2023

No âmbito da inspeção dimensional, é uma prática pouco comum aproveitar as áreas periféricas dos sensores devido aos possíveis efeitos de distorção introduzidos pelas lentes. Essa distorção, que assume um caráter não linear, apresenta desafios na sua modelagem. No entanto, quando abordada com detalhamento, abre caminho para uma utilização otimizada dos sensores ópticos – componentes cujos custos frequentemente se elevam proporcionalmente à resolução.

Os benefícios desse enfoque são significativos. Primeiramente, porque permite utilizar uma região maior, se não total, do campo de visão da câmera. Dessa forma se torna possível a inspeção simultânea de múltiplas características dimensionais de uma peça, que não caberiam na região de distorções desprezíveis da imagem, em uma única captura, sem demandar reposicionamento da peça ou da câmera. Isto, portanto, se traduz em ganho significativo de eficiência, ao diminuir o número de aquisições necessárias na inspeção. Não somente, utilizar as áreas periféricas do sensor proporciona um aproveitamento mais otimizado dos pixels disponíveis. Em sistemas de inspeção nos quais o valor da câmera

guarda uma relação direta com a resolução, essa abordagem contribui positivamente com a diminuição do custo, pois permite a utilização de equipamentos mais acessíveis para satisfazer os requisitos de campo de visão.

Embora seja preciso reconhecer a complexidade que envolve a calibração para compensar as distorções, avanços em técnicas de modelagem e algoritmos de *machine learning* têm simplificado processos e agilizado o desenvolvimento de software (GOODFELLOW et al., 2016). Dessa forma, avaliar uma técnica de calibração com *machine learning* que considere as áreas periféricas dos sensores ópticos para inspeção dimensional é um estudo pertinente. Pois, ao superar as limitações das distorções através da calibração, é possível colher vantagens notáveis em termos de eficiência e redução de custos, desde que a depreciação da incerteza de medição não seja significativa. Diante da contínua evolução das tecnologias de modelagem e inspeção, espera-se que a abordagem de IA ganhe ainda mais relevância, oferecendo soluções práticas e eficazes para os desafios presentes nesse campo.

1.1 Objetivos

Com a ambição de ampliar a área do sensor utilizada para tarefas de inspeção dimensional de alta precisão, este trabalho propõe um método de calibração de sistema de visão com câmera de linha, sensível a efeitos não lineares de distorção da lente. Além disso, ensaio prático em ambiente real.

O ambiente de captura será definido pela instalação dos equipamentos disponíveis, considerando suas características técnicas. Serão descritos os trabalhos de projeto e instalação do sistema de visão e, posteriormente, de projeto e confecção do padrão de calibração utilizado.

A validação do método proposto será realizada a partir de ensaio prático, em termos de erro de reprojeção. A partir de uma imagem do padrão serão extraídas as coordenadas em pixels dos pontos de referência, cujas coordenadas reais da peça física, em milímetros, são conhecidas.

O objetivo específico será calibrar o sistema para realização de inspeção dimensional com tolerância de 1 mm.

1.2 Estrutura

Este documento está organizado em cinco capítulos, incluindo esta introdução na qual foi apresentado o contexto tecnológico no qual está inserido o trabalho, bem como enunciados seus objetivos específicos. No próximo, Capítulo 2, será detalhada a revisão da literatura e estudos decorrentes realizados para fundamentar a abordagem selecionada para solucionar o problema descrito.

A seguir, no Capítulo 3, são descritos os equipamentos, parâmetros e métodos utilizados, detalhados de modo a permitir sua reprodução a partir deste documento. Posteriormente, o Capítulo 4 reúne os resultados obtidos e sua discussão, com os aprendizados do trabalho. Finalmente, no Capítulo 5 são sumarizadas as principais contribuições desta pesquisa, suas limitações e oportunidades de desenvolvimento adicional.

2 REVISÃO BIBLIOGRÁFICA E ESTADO DA ARTE

A visão computacional é um campo interdisciplinar que combina técnicas de processamento de imagens, aprendizado de máquina e inteligência artificial para permitir que computadores entendam e interpretem o mundo visual de maneira semelhante aos seres humanos (GONZALEZ et al., 2010). Ela se concentra na obtenção, processamento e análise de informações a partir de imagens e vídeos. A visão computacional tem diversas aplicações, incluindo reconhecimento de objetos, detecção de padrões, segmentação de imagens, rastreamento de objetos, realidade aumentada, medicina e automação industrial (GONZALEZ et al., 2010).

Ao longo dos anos, houve avanços significativos na área de visão computacional, impulsionados por melhorias na capacidade de processamento, na disponibilidade de conjuntos de dados grandes e rotulados, e no desenvolvimento de algoritmos mais sofisticados. Alguns dos principais tópicos e técnicas dentro da visão computacional incluem, (GONZALEZ et al., 2010), (JABŁOŃSKI et al., 2016):

- **Processamento de Imagens:** O processamento de imagens envolve a aplicação de operações matemáticas e algoritmos para melhorar, realçar ou modificar imagens de modo a facilitar a análise posterior. Isso inclui filtragem, segmentação, restauração e transformações geométricas.
- **Detecção e Rastreamento de Objetos:** Essas técnicas visam localizar e acompanhar objetos específicos em uma sequência de imagens ou em um vídeo. Algoritmos de detecção podem usar abordagens baseadas em características, como histogramas de cores e texturas, ou em aprendizado de máquina, como redes neurais convolucionais (CNNs).
- **Reconhecimento de Padrões e Classificação:** O reconhecimento de padrões envolve identificar objetos ou padrões em uma imagem e atribuir rótulos a eles. Isso é frequentemente alcançado usando técnicas de aprendizado supervisionado ou não supervisionado, como redes neurais, máquinas de vetores de suporte (SVMs) e outros métodos de classificação.
- **Segmentação de Imagens:** A segmentação visa dividir uma imagem em regiões significativas ou objetos individuais. Isso é essencial para a análise detalhada de elementos específicos em uma cena.

- Visão 3D e Realidade Aumentada: Essas áreas exploram a reconstrução tridimensional de cenas a partir de imagens 2D, permitindo a sobreposição de informações virtuais no mundo real.
- Aprendizado de Máquina e Redes Neurais: O aprendizado de máquina, especialmente as redes neurais, revolucionou a visão computacional, permitindo a criação de modelos que podem aprender representações hierárquicas complexas diretamente dos dados.
- Processamento de Vídeo: Além de imagens estáticas, a visão computacional também lida com a análise de vídeos, incluindo rastreamento de objetos em movimento, detecção de atividades e reconhecimento de padrões temporais.

No geral, a visão computacional desempenha um papel crucial em várias indústrias e campos de pesquisa, desde a automação industrial até a medicina e a tecnologia de veículos autônomos. É um campo em constante evolução, com novas abordagens e técnicas sendo desenvolvidas regularmente para lidar com os desafios da interpretação computacional do mundo visual.

2.1 Câmeras de linha

As câmeras de linha, também denominadas câmeras lineares, possuem um papel significativo na área de visão computacional, oferecendo uma abordagem distintiva para aquisição de imagens em cenários específicos. Diferentemente das tradicionais câmeras de matriz, que capturam a cena inteira, as câmeras de linha registram sequencialmente uma única linha de pixels, tornando-as especialmente vantajosas para aplicações como inspeção industrial (LEE et al., 2002), leitura de códigos de barras e digitalização de documentos.

Esses dispositivos são compostos por uma matriz unidimensional de sensores de imagem, organizados de forma sequencial. Através de um sistema óptico, a luz da cena é direcionada para esses sensores, seja movendo a câmera ou a cena em relação a ela, resultando na aquisição sequencial das diferentes linhas da cena.

As vantagens técnicas das câmeras de linha são notáveis. Sua capacidade de capturar imagens linha por linha agiliza o processo, sendo particularmente adequadas para situações de alta velocidade. São um tipo de sensor capaz de reduzir problemas decorrentes do borrão de movimento, do inglês *blur*, tornando-as valiosas para aplicações industriais que exigem detalhes nítidos de objetos em deslocamento. Dessa forma, são especialmente interessantes também para uso em tecnologias espaciais (GUPTA et al., 1997).

No domínio da inspeção industrial, as câmeras de linha desempenham um papel de destaque, detectando defeitos em produtos em movimento, tais como rachaduras e imperfeições de superfície. Em contextos de automação, elas são empregadas para garantir a qualidade, contribuindo para uma produção eficiente. Entretanto, há desafios técnicos a serem enfrentados. A calibração precisa da geometria da câmera e do sistema óptico é importante para determinar a incerteza de medição do sistema. Além disso, é fundamental

sincronizar meticulosamente o movimento do objeto ou da câmera com a captura da linha, evitando artefatos indesejados na imagem.

Em suma, as câmeras de linha ocupam um lugar de destaque na visão computacional, proporcionando uma abordagem única em situações onde a velocidade de aquisição e a minimização do borrão de movimento são requisitos. Seu papel na inspeção industrial, automação e controle de qualidade é incontestável, destacando-se como uma solução tecnológica de grande valor.

2.2 Calibração de sistemas de visão

A calibração de sistemas de visão se faz necessária, entre outras, em aplicações de dimensionamento de precisão, nas quais o objetivo é extrair da imagem as dimensões reais do objeto planejado, como nos trabalhos de Lee et al. (2002), Lutz et al. (1994), C.A. Luna et al. (2006) e Petty et al. (1998). O mesmo se aplica para tecnologias de digitalização a partir de imagens, pois a identificação do sistema pelo processo de calibração permite remover distorções da imagem e obter a descrição de um plano real, físico.

Esta identificação consiste em encontrar uma relação determinística entre coordenadas reais, em milímetros por exemplo, e coordenadas da imagem, em pixels (HORAUD et al., 1993). Ou seja, consiste na identificação de um modelo matemático que descreva a fotografia. A complexidade deste modelo pode assumir diversos graus, de acordo com a configuração do ambiente de inspeção e incerteza tolerada, diretamente relacionada à precisão requerida.

Além disso, a calibração de câmeras de linha é relevante para diversas outras aplicações, pois permite corrigir, ou reduzir, imprecisões relacionadas a instalação da câmera (DI LEO et al., 2017). Por exemplo, desalinhamento entre direção de movimento e sensor da câmera, que devem ser perpendiculares para construção de uma imagem precisa. Caso contrário, a imagem obtida terá pixels com distorção *skew*¹, (STEGER; ULRICH; WIEDEMANN, 2018).

Cabe destacar que sincronia imperfeita entre velocidade de movimento do objeto e velocidade de disparo da câmera pode causar efeitos de esticamento ou encolhimento do objeto capturado. Ainda que suaves, estes efeitos podem ser ampliados por operações computacionais de processamento de imagem, ou mesmo comprometer os resultados. Pode-se citar segmentação, extração de *features* (processamento de momentos ou *features* de forma de regiões segmentadas), semelhança de modelos 2D, semelhança estéreo. A calibração do sistema de visão permite corrigir as imagens reduzindo os efeitos de distorção e *skew* para níveis desprezíveis, virtualmente os eliminando. Dessa forma, simplifica a instalação da câmera, uma vez que torna desnecessário um alinhamento preciso (STEGER; ULRICH, 2021).

¹Fenômeno de distorção transitória visível em imagens capturadas por câmeras que empregam determinados tipos de sensores CMOS, manifestando-se como uma inclinação nas linhas verticais de pixels

2.2.1 Métodos de calibração

As abordagens de calibração podem ser divididas com base nos modelos de câmera de linha utilizados: estático ou dinâmico. Enquanto na modelagem estática, o movimento da câmera não é considerado no equacionamento, modelos dinâmicos permitem calibrar o sistema incluindo o avanço da câmera ou da cena (STEGGER; ULRICH, 2021).

Na abordagem de calibração de modelos estáticos, diversas proposições de método e padrões associados já foram estudadas: Horaud et al. (1993) propôs a utilização de *cross-ratio invariance* para fins de calibração de câmeras de linha, porém requeria deslocamentos precisos do padrão utilizado; mais tarde, Carlos A. Luna et al. (2010) apresentou resultados de calibração a partir de um mesmo padrão repetido em planos paralelos e diferentes alturas, sem necessidade de quaisquer incrementos posicionais; e recentemente, Liao et al. (2019) avaliou o erro de excentricidade em padrões de linhas sólidas.

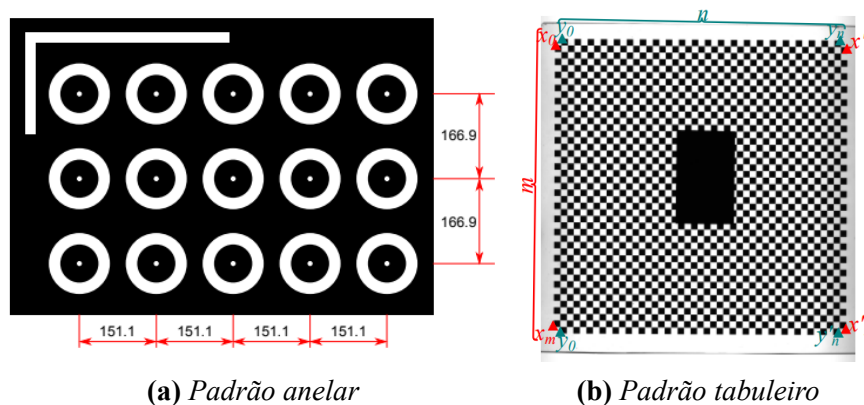
Modelos dinâmicos, por sua vez, requerem calibração por escaneamento, que consiste na aquisição de padrões planares, por meio do movimento relativo entre alvo e câmera, de modo a construir uma representação bidimensional. Esta abordagem se assemelha às técnicas de calibração de câmeras convencionais, no que diz respeito a identificação de pontos de referência e refinamento dessas coordenadas. Oferece também a praticidade de permitir utilizar de padrões já existentes e vastamente empregados em calibração, como padrões de tabuleiro (ZHANG et al., 2018) e grades de círculos (WENDEL et al., 2017).

2.2.2 Padrões de calibração

Padrões de calibração, no contexto de visão computacional, são compostos por um conjunto de pontos de referência de distribuição conhecida. Possuem geometrias simples, com objetivo de caracterizá-los bem na imagem, de modo a facilitar a identificação das coordenadas de projeção para cada ponto (DUDA et al., 2018).

Os mais comuns são os padrões do tipo anelar (WENDEL et al., 2017) e tipo tabuleiro (ZHANG et al., 2018), ilustrados na Figura 2.

Figura 2: Exemplos de padrão de calibração de sistemas de visão.



Fonte: Wendel et al. (2017), Zhang et al. (2018)

Niu et al. (2018), Heikkila e Silven (1997) e Heikkila (2000) se utilizaram de padrões anelares em seus trabalhos, que são mais robustos em relação a variações de distância e

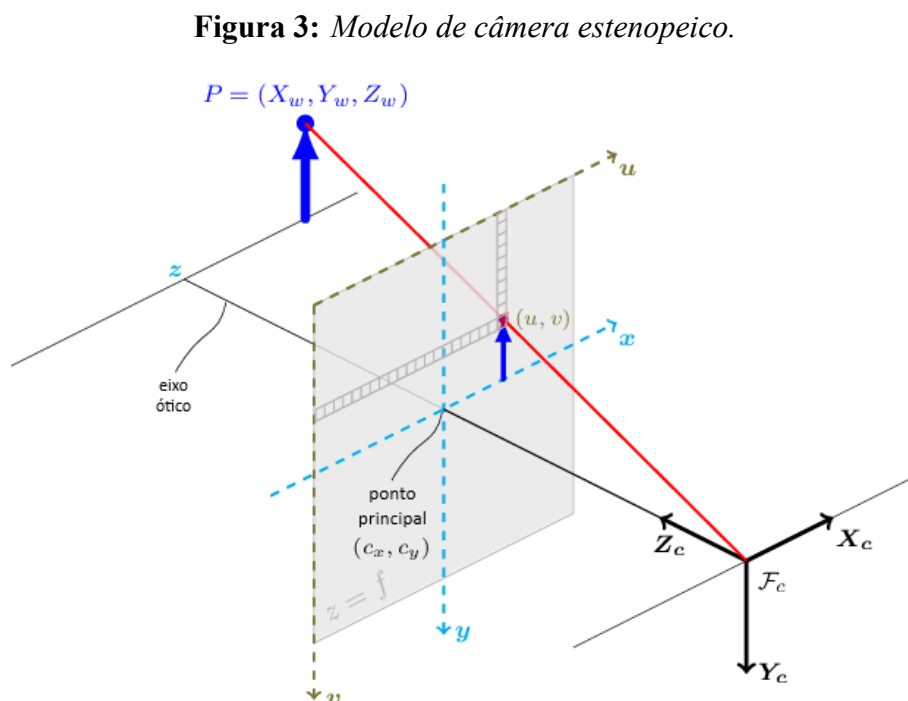
angulação da câmera, em termos de precisão dos pontos de referência. Porém sua detecção é mais complexa, ou seja, o custo computacional para identificação dos pontos é mais alto que o de padrões do tipo tabuleiro, e sua eficiência espacial é reduzida, uma vez que cada círculo ocupa um espaço significativo da imagem.

Alternativamente, Donné et al. (2017) e Ying et al. (2014) reportaram resultados com padrões tabulares, que se destacam por permitir alta densidade de pontos, quando comparados com as alternativas, e facilidade de detecção dos pontos na imagem. Em sistemas em que a instalação da câmera é fixa e a distância do objeto é constante, como a estudada neste trabalho, são especialmente interessantes, pois são mais sensíveis à alterações de distância e ângulo.

2.3 Modelo da câmera de linha

Neste sistema de visão, devido à utilização de câmeras de linha, a largura da imagem (eixo x) é definida pela linha de captura, sendo do tamanho do sensor; diferentemente, a altura da imagem (eixo y) é definida pelo empilhamento de capturas sequenciais de acordo com o movimento do objeto sobre a esteira. Dessa forma é possível utilizar câmeras 1D para geração de imagens 2D.

O modelo estenopeico descreve o funcionamento de câmeras por projeção perspectiva. Ele considera que todos os raios de luz se intersectam em um mesmo ponto e, a partir disso, permite mapear geometricamente posições 3D para o plano da imagem 2D. A Figura 3 ilustra, evidenciando o centro da câmera C , o eixo principal Z , bem como o ponto principal p .



Fonte: OpenCV, acessado em 2023

O centro da câmera C , também chamado de centro óptico, é o ponto de convergência dos raios de luz e origem do sistema de coordenadas. Por ele passa o eixo Z , perpendi-

cular ao sensor da câmara, denominado por eixo principal ou eixo óptico. O plano $Z = f$ descreve o plano da imagem, ou plano focal, uma vez que f expressa a distância focal da câmara. Seu ponto de intersecção com o eixo óptico é chamado de ponto principal, da imagem p .

Dessa forma, um ponto $(X, Y, Z)^T$ no espaço de coordenadas XYZ é mapeado para o plano da imagem uv por:

$$(u, v)^T = \left(\frac{fX}{Z}, \frac{fY}{Z} \right)^T \quad (2.1)$$

Entretanto, a origem das coordenadas dos sistemas de imagens é comumente o pixel mais à esquerda e mais acima da imagem e, portanto, requer uma conversão do sistema de coordenadas. Para tanto, utiliza-se as coordenadas do ponto principal. Em coordenadas homogêneas:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (2.2)$$

A equação (2.2) é válida para imagens que possuem a mesma escala nas direções dos eixos x e y , ou seja, os pixels têm formato quadrado. Para alguns tipos de câmara essa premissa não é aplicável, incluindo a câmara de linha pois os parâmetros que definem a escala na direção do movimento são independentes dos que definem a escala na direção do sensor de linha. Em função disso, o modelo estenopeico não é aplicável ao longo do eixo y , que depende apenas dos parâmetros frequência de disparo e velocidade da esteira. Na equação (2.3), tais fenômenos são expressos pela variável k_y .

É possível considerar ainda outros fatores, como *skew* e as distorções da lente. O *skew* é conhecido em câmeras 2D como uma inclinação do sensor. Ocorre, em câmeras de linha, quando a direção de movimento do objeto e o sensor da câmara não são perpendiculares entre si, causando distorção na projeção obtida. Esse efeito pode ser descrito pela inclusão do parâmetro s no modelo, conforme expresso na equação (2.3). Já as distorções da lente, matéria de interesse deste estudo, serão descritas com maior detalhamento na sequência.

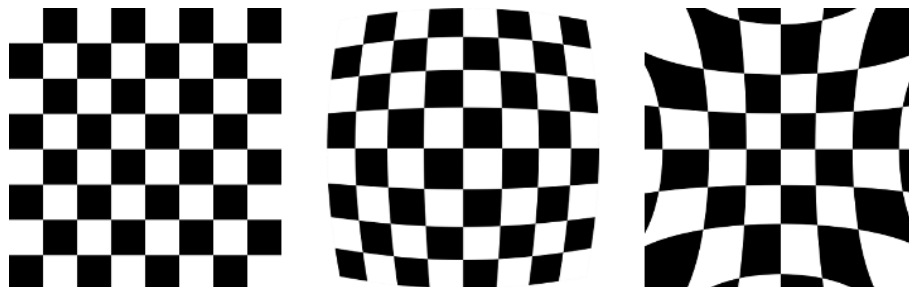
$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & s & p_x & 0 \\ 0 & k_y & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (2.3)$$

2.3.1 Distorção da lente

As lentes reais possuem imperfeições e, por isso, propagam distorções para a imagem. A principal delas, de acordo com Zisserman (2004) é a distorção radial, cujo efeito aumenta com a distância do ponto principal, conforme ilustrado na Figura 4. À esquerda, uma representação de um tabuleiro sem distorções, no centro o efeito de distorção radial negativa, também chamada de distorção do tipo barril e, à direita, distorção radial positiva.

Existe também a distorção tangencial, causada pelo não paralelismo entre a lente e o sensor da câmera.

Figura 4: *Representação gráfica da distorção radial.*



Fonte: OpenCV, acessado em 2023

Conforme discutido anteriormente, pode-se utilizar esse modelo somente para descrição do eixo x da imagem, paralelo ao sensor da câmera. Os parâmetros intrínsecos são os parâmetros utilizados para descrever a relação entre as coordenadas da câmera, cuja origem está posicionada no centro óptico, e as coordenadas da imagem.

2.4 Aprendizado de máquina

A área do Aprendizado de Máquina (ML, do inglês *Machine Learning*) é um campo da Inteligência Artificial que se dedica a empregar algoritmos para solucionar problemas práticos por meio da análise de conjuntos de dados e do treinamento de modelos estatísticos (BURKOV, 2020). A base do ML reside na capacidade dos sistemas computacionais aprenderem e melhorarem automaticamente seu desempenho, sem a necessidade de programação direta para cada tarefa. De modo geral, o ML assume um papel de destaque para lidar com problemas de elevada complexidade, que exigem uma compreensão profunda do ambiente físico. Ainda que intuitivas para humanos, como pegar uma bola de tênis em movimento, a tradução desse tipo de conhecimento para um computador representa um desafio significativo pela perspectiva analítica. Goodfellow et al. (2016) destacam uma variedade de problemas solucionados por meio do Aprendizado de Máquina, enfatizando:

- **Regressão:** Nesse âmbito, algoritmos são treinados para prever valores numéricos com base em dados de entrada. Um exemplo prático é a avaliação de crédito, onde um modelo pode estimar o valor de crédito apropriado a ser concedido a um cliente com base em históricos de dados.
- **Classificação:** Embora não seja o foco deste estudo, a classificação desempenha um papel crucial no Aprendizado de Máquina. Nesse cenário, algoritmos categorizam novos dados de entrada em uma entre várias categorias predefinidas. Um exemplo é o reconhecimento de objetos em imagens, onde o sistema atribui a categoria correta a um objeto com base em suas características visuais.
- **Detecção de Anomalias:** Outra área relevante é a detecção de padrões anômalos em sequências de eventos. Ela é frequentemente empregada na monitorização de má-

quinas industriais, onde variações anormais nas vibrações dos componentes podem indicar falhas futuras.

Além disso, os algoritmos de ML podem ser categorizados por método, conforme descrito por Russel et al. (2020), em três grupos principais: aprendizado não-supervisionado, em que os modelos identificam padrões nos dados de entrada sem requerer exemplos de saída esperada; aprendizado supervisionado: nesse caso, os modelos aprendem por meio de exemplos de pares de entrada e saída desejada, ajustando-se para fazer previsões precisas em novos dados de entrada; e aprendizado por reforço: no qual os modelos aprendem por meio de recompensas e penalidades baseadas em suas ações, frequentemente aplicadas a problemas de tomada de decisão sequencial.

O uso de ML no procedimento de calibração de sistemas de visão com câmeras de linha pode ser justificado pela sua praticidade, concedendo agilidade de desenvolvimento. Além disso, a capacidade de aprendizado sensível a fenômenos não lineares permite explorar aplicabilidade de técnicas nesse contexto. A calibração de sistemas de visão, especialmente com câmeras de linha, envolve uma série de desafios, como a modelagem das características geométricas da câmera, a correção de distorções e a adequação ao ambiente de operação. Abordagens tradicionais muitas vezes podem não capturar todos esses aspectos de maneira eficaz, resultando em calibrações imprecisas e ineficientes.

Machine Learning, especialmente técnicas como redes neurais, pode aprender a modelar as não linearidades, distorções e variações complexas do sistema de câmera, permitindo uma calibração mais precisa e abrangente. Essas técnicas podem lidar com as variações ambientais, como mudanças na iluminação e condições de operação, adaptando-se automaticamente para obter resultados confiáveis em cenários do mundo real.

Além disso, a automação é um dos pontos fortes do ML. Enquanto a calibração tradicional frequentemente requer intervenção manual e várias iterações, o *Machine Learning* pode automatizar significativamente o processo. Isso não apenas reduz a intervenção humana, mas também acelera o tempo de calibração, tornando o processo mais eficiente.

Outra vantagem é a capacidade de aprendizado contínuo. À medida que o sistema de visão é usado e coleta mais dados ao longo do tempo, o ML pode adaptar gradualmente o modelo de calibração, mantendo a precisão mesmo em situações em constante evolução. Isso é particularmente valioso em ambientes de produção que estão sujeitos a mudanças graduais ao longo do tempo.

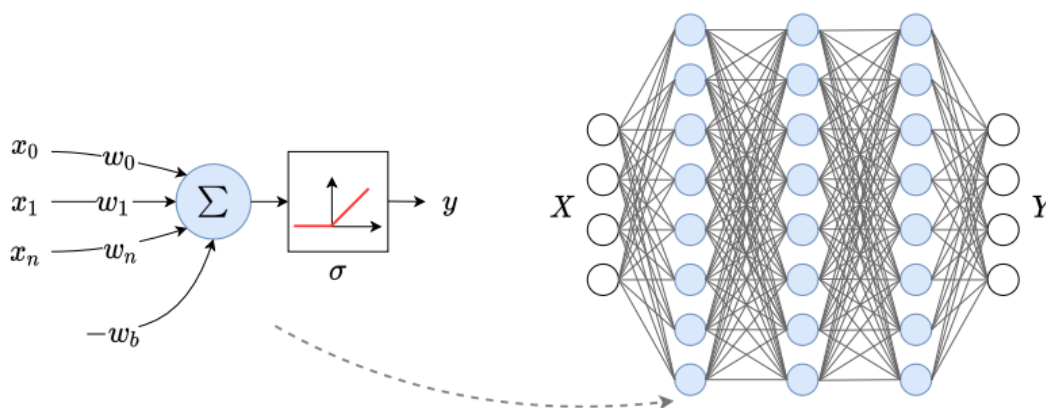
Em última análise, a incorporação de *Machine Learning* no procedimento de calibração de sistemas de visão com câmeras de linha oferece uma abordagem mais adaptativa, eficiente e precisa para lidar com a complexidade e as variações presentes no processo de calibração. Oferece, portanto, resultados mais confiáveis, redução de erros humanos e melhoria contínua da precisão à medida que o sistema é usado ao longo do tempo.

2.4.1 Redes neurais

O perceptron, introduzido em 1957 por Frank Rosenblatt, (ROSENBLATT, 1958), é um elemento essencial no desenvolvimento das redes neurais e do aprendizado profundo. Inicialmente usado para resolver problemas de classificação linear, seu verdadeiro potencial surgiu com as redes neurais multicamadas, como as MLPs. Essas redes permitiram

a superação das limitações iniciais ao capturar relações não-lineares. A Figura 5 ilustra a estrutura do perceptron. À esquerda, o modelo do neurônio artificial *Perceptron*, com entradas x_i , pesos w_i , viés, do inglês *bias*, w_b , função de ativação σ e saída y . À direita, uma rede neural profunda com três camadas completamente conectadas.

Figura 5: Modelo do neurônio artificial *Perceptron* (esquerda) e uma rede neural profunda com três camadas totalmente conectadas (direita).



Fonte: Rosa (2021)

O advento do aprendizado profundo trouxe destaque às redes neurais convolucionais (CNNs) e recorrentes (RNNs). Ao longo dos anos, essas arquiteturas evoluíram, possibilitando a resolução de tarefas complexas em visão computacional, processamento de linguagem natural e muito mais. O perceptron, fundamental na base dessas redes, desempenha um papel crucial no processamento de informações e extração de padrões de dados.

O *Multi-Layer Perceptron* (MLP) é uma rede neural flexível com camadas intermediárias, eficaz para resolver desafios que envolvem relações complexas entre dados. É frequentemente aplicado em tarefas como classificação, regressão e análise de sentimentos em textos, onde a interconexão de características é crucial (CHAI et al., 2021).

As *Convolutional Neural Networks* (CNNs), por outro lado, são estruturadas para lidar especificamente com dados como imagens e vídeos. Essas redes são especialmente úteis em problemas de visão computacional, como classificação de imagens e detecção de objetos, bem como em domínios como processamento de áudio e análise de séries temporais, onde a estrutura dos dados se assemelha à de imagens (VOULODIMOS et al., 2018).

As *Recurrent Neural Networks* (RNNs) destacam-se quando os dados são sequenciais, como em linguagem natural e séries temporais. São empregadas em tarefas como previsão de séries temporais, tradução automática e geração de texto. Variantes avançadas, como as *Long Short-Term Memory* (LSTM) e *Gated Recurrent Unit* (GRU), atenuam as limitações de degradação do gradiente em sequências longas (CHAI et al., 2021).

Neste trabalho será utilizada uma rede MLP, dada sua capacidade de generalização de características não-lineares, aliada à simplicidade dos dados que não requerem atributos sofisticados de visão computacional.

3 MATERIAIS E MÉTODOS

Neste capítulo foram reunidas as especificações técnicas dos equipamentos utilizados no trabalho, projeto do padrão de referência, características de montagem do sistema de visão, roteiro do ensaio de calibração e modelagem do algoritmo de IA proposto.

Os estudos foram realizados em um sistema de visão composto por câmera de linha, encoder, plataforma de deslocamento linear, iluminador e um computador. A montagem do sistema consistiu em posicionar a câmera de linha acima da plataforma que transporta objetos. O iluminador foi colocado ao lado da câmera de linha para iluminar o objeto em movimento. A câmera de linha foi conectada ao computador que processou as imagens capturadas pela câmera de linha, para visualização e exportação. O encoder, acoplado ao eixo do mecanismo de movimento linear, foi conectado diretamente à câmera de linha e utilizado como sinal de disparo. A câmera foi configurada para que a frequência de aquisição gerasse imagens com resolução em v , direção paralela ao sentido de avanço do objeto, similar à resolução em u , com capturas disparadas a partir do sinal do encoder.

As especificações técnicas dos equipamentos disponíveis foram consideradas para posterior escolha dos parâmetros de ensaio.

A câmera utilizada foi uma Basler racer GigE, modelo raL4096-24gm, com resolução de 4096 px, monocromática, tamanho de pixel $7\mu m \times 7\mu m$, frequência máxima de disparo 26 kHz, tempo de exposição programável, alimentação de 12 a 24 VDC e consumo de 4 W. Com suporte para saída de dados em Fast Ethernet (100 Mbit/s) ou Gigabit Ethernet (1000 Mbit/s) e montagem compatível com tipos F, C, M42x1 e M42x0.75. Para configuração do dispositivo, bem como aquisição de imagens, foi utilizado o software Basler pylon Camera Software Suite (v5), no Windows 7. A lente utilizada foi uma Azure-2428MT, de montagem M42x1, distância focal de 24 mm, faixa de abertura de F2.8 a F22, ângulo de abertura de 84° e foco mínimo de 0,15 m.

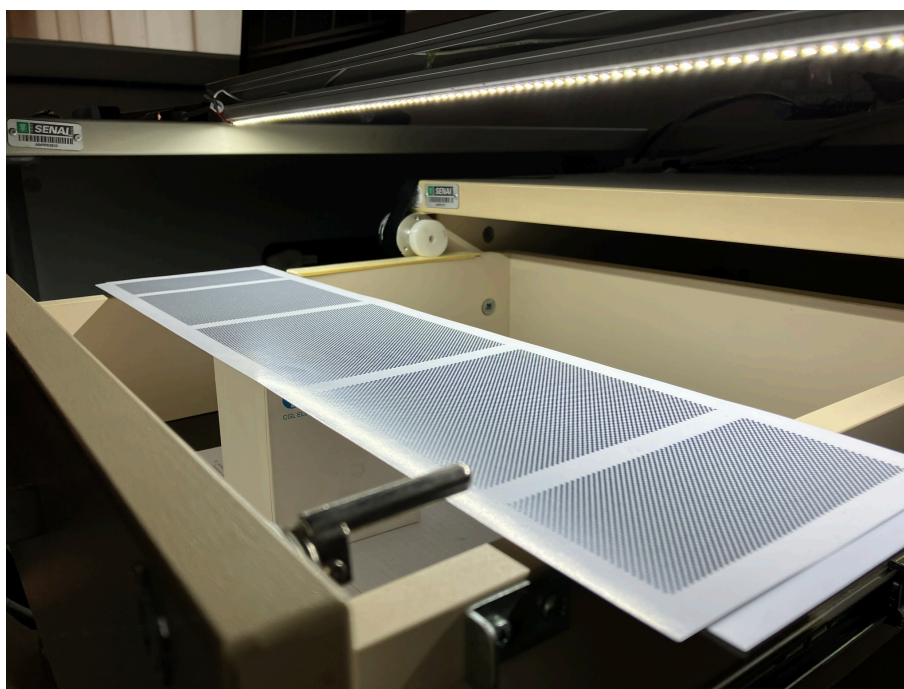
O movimento entre câmera e objeto foi realizado por meio de um mecanismo de deslocamento linear de acionamento manual, Figura 6. Composto por uma plataforma sobre um par de corredeiras telescópicas paralelas, com extensão máxima de 0,25 m. A utilização de um *encoder* no sistema de visão tem por objetivo sincronizar o disparo das imagens com a velocidade de movimento da esteira, ou elemento móvel. Dessa forma torna o sistema mais robusto, independente de variações na velocidade da esteira.

O *encoder* utilizado foi um E6B2-CWZ5B, do tipo incremental, 2.000 pulsos por rotação, alimentação de 12 a 24 VDC, com configuração de saída coletor aberto PNP. O acoplamento à superfície de deslocamento linear foi por roda emborrachada de 0,3 m.

O registro de objetos em movimento com nitidez requer tempos de exposição tão pequenos quanto possível, de modo a minimizar o efeito de *blur* deixado na imagem. Isto acontece porque a linha de captura está em constante deslocamento sobre o objeto observado. Idealmente, quanto menor o tempo de exposição maior a nitidez da imagem, desconsiderando-se demais fatores.

Entretanto deve-se considerar a condição de iluminação do ambiente de captura, pois com a diminuição do tempo de exposição é reduzida a resposta do sensor, uma vez que menos luz é recebido por este. Lâmpadas típicas, com alimentação de corrente alternada, emitem luz de forma oscilatória com frequência de 60 Hz, frequência da rede elétrica geral. Apesar deste efeito não ser perceptível a olho nu, pode ser registrado por câmeras fotográficas como a utilizada neste trabalho. Uma vez que a intensidade luminosa varia ao longo do tempo e, o tempo de exposição utilizado é suficientemente pequeno, cada captura, ou seja, cada linha possui condição específica de iluminação, deixando a imagem 2D com um efeito de inúmeras listras horizontais bastante evidente, similar à sombra de uma persiana laminar entreaberta. De modo a evitar este problema, o iluminador utilizado foi um filamento de LED, de 0,4 m sem difusor, com alimentação DC.

Figura 6: Padrão de calibração sobre base de deslocamento linear.



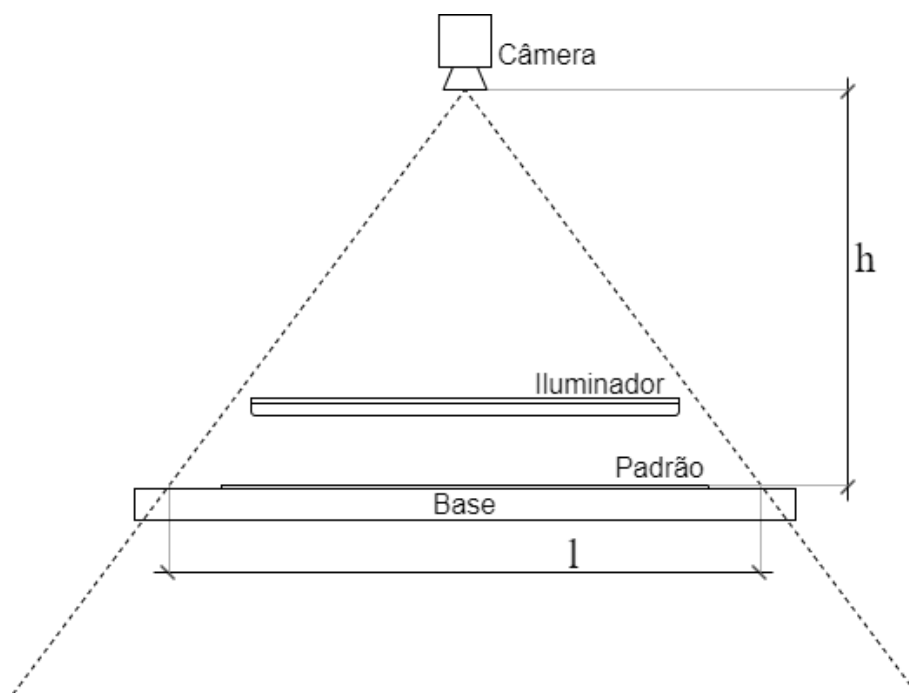
Fonte: autor (2022)

O esquemático da Figura 7 descreve o sistema de visão na configuração utilizado para o ensaio de calibração proposto, tal que $l = 0,45$ m e $h = 0,50$ m.

3.1 Padrão de referência

O projeto do padrão de calibração foi desenvolvido a partir de um ensaio exploratório estático, no qual foram identificadas características de nitidez da imagem ao longo de toda a linha de captura, em condições de operação similares ao ensaio de calibração.

Figura 7: Sistema real de captura.

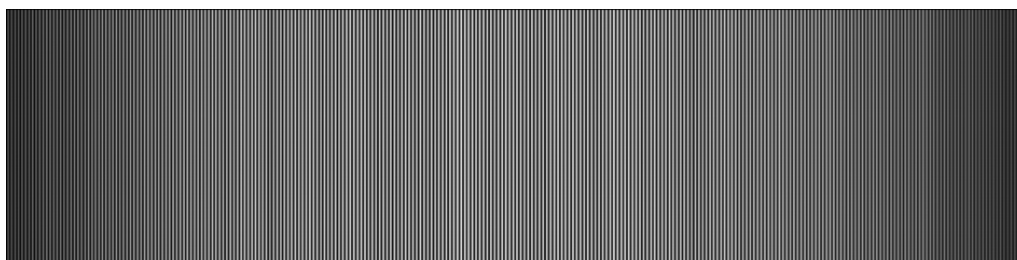


Fonte: autor (2023)

A câmera foi posicionada a aproximadamente 0,6 m acima do chão, onde estava disposta uma impressão de listras brancas e pretas em sequência, cada uma com 1 mm de largura. O comprimento total dessa peça era de 1,4 m, excedendo o campo de visão de 0,55 m.

Uma sequência de 1024 capturas foi realizada, resultando em imagens 2D conforme Figura 8, em diversas condições de abertura do obturador e tempo de exposição.

Figura 8: Imagem obtida no ensaio estático.



Fonte: autor (2022)

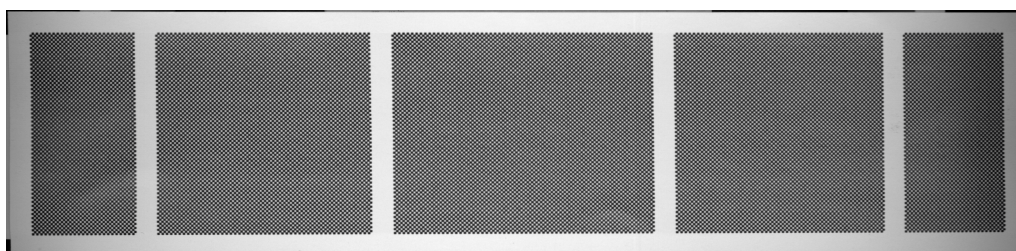
A de melhor contraste, sem apresentação de saturação, foi selecionada para análise, na qual foram identificados os comprimentos das faixas de transição entre branco e preto, de 5 a 7 pixels. A partir disso, foi selecionada a dimensão das casas do padrão de calibração do tipo tabuleiro: 1 mm x 1 mm.

O padrão proposto foi impresso em papel couchê, composto por cinco tabuleiros preto e branco lado a lado, com distanciamento de 10 mm entre si e casas de 1 mm x 1 mm, Figura 9. Devido a limitação do processo de impressão, considerando-se alternativas de baixo custo, as dimensões totais da peça foram de 100x460 mm, sem considerar as bordas

brancas em torno dos tabuleiros, que são necessárias para que o algoritmo de detecção dos pontos funcione corretamente.

Os tabuleiros periféricos, identificados por 1 e 5, são compostos por 100 casas de altura e 50 de largura, cada um com 4.851 pontos de referência. O central, de número 3, de mesma altura e 120 casas de largura, possui 11.781 pontos, enquanto os intermediários, tabuleiros 2 e 4, de 100x100 casas, possuem 9.801 cada. No total, o padrão de calibração possui 41.085 coordenadas.

Figura 9: Imagem do padrão de calibração composto pelos tabuleiros 1 a 5, da esquerda para a direita.



Fonte: autor (2022)

3.2 Ensaio de calibração

O procedimento de calibração proposto consistiu na realização dos seguinte roteiro:

1. Instalação da câmera 0,45 m acima da plataforma de deslocamento linear, para capturar toda a extensão do padrão. Para tanto, foi utilizado um tripé telescópico.
2. Ajuste da posição para garantir alinhamento perpendicular entre o sensor da câmera e a direção de movimento da base móvel. Como referência, foi utilizada a borda da plataforma.
3. Configuração do tempo de exposição em 1 ms.
4. Aumento da abertura do obturador para ajuste fino da imagem com finalidade de:
 - (a) evitar saturação; e
 - (b) melhorar a nitidez das bordas das casas do tabuleiro.
5. Por fim, captura de imagem do padrão de calibração.

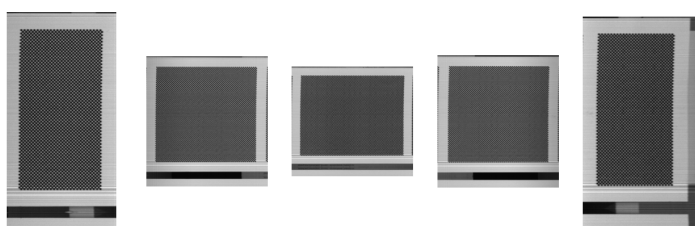
3.3 Sistema de visão

O método de calibração proposto requer uma única imagem, visto que a posição da câmera é fixa e sua distância do objeto, sobre o mecanismo de deslocamento, constante. Neste trabalho, foram realizadas 10 aquisições para avaliar o padrão projetado em termos de reconhecimento dos pontos de referência, conforme discutido adiante, Seção 4.1.

3.3.1 Preparo dos dados

O software foi desenvolvido em Python 3, em Jupyter Notebook, com utilização das bibliotecas NumPy (HARRIS et al., 2020), Matplotlib (HUNTER, 2007), Pandas (MCKINNEY et al., 2010), OpenCV (BRADSKI, 2000) e Scikit-learn (PEDREGOSA et al., 2011). Na primeira etapa, a imagem do padrão de calibração é segmentada em cinco regiões, com parâmetros ajustados manualmente para o ensaio realizado. Cada segmento contém um dos cinco tabuleiros componentes do padrão fabricado, conforme ilustrado na Figura 10.

Figura 10: Segmentação dos tabuleiros 1 a 5 em imagens separadas.



Fonte: autor (2022)

Após, é realizada a busca dos pontos de referência de cada segmento, por meio da função *findChessboardCornerSB*, que retorna as posições dos cantos interiores de um tabuleiro usando a abordagem proposta por Duda et al. (2018). A função requer informações de tamanho do tabuleiro para identificar corretamente seus cantos internos e, além disso, pressupõe representação clara de todas as casas. Ou seja, requer uma borda lisa e imagem sem oclusão de qualquer região da malha.

Os dados, que consistem nas coordenadas dos cantos internos dos tabuleiros, são então concatenados e reordenados, para que estejam consistentes com os dados de referência. Estes, medidos em ensaio de calibração tridimensional, no laboratório de metrologia de ISI-SIM, Figura 11.

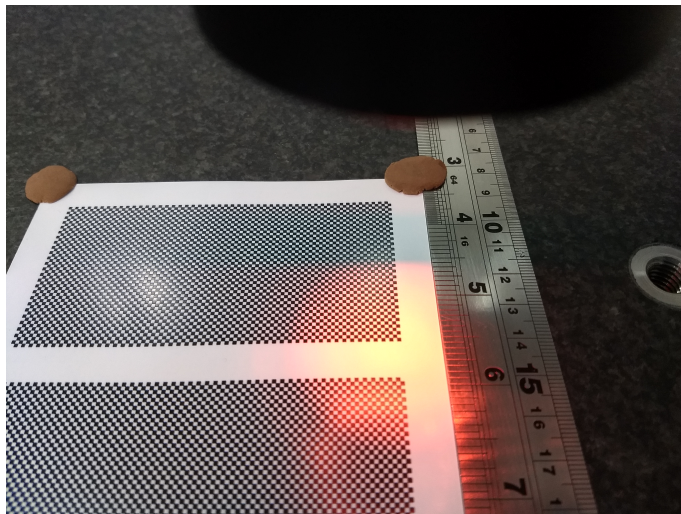
3.3.2 Algoritmo de IA

De posse dos 41.085 pares de dados reais e projetados, respectivamente sem distorção e distorcidos, foi treinada uma rede de IA MLP para regressão. Os agrupamentos de treinamento e teste foram selecionados aleatoriamente do conjunto de dados, com a função *train_test_split* da biblioteca *scikit-learn*, na proporção de 70% e 30%. Ou seja, 28.759 pontos de referência foram utilizados na etapa de aprendizado, de treino, enquanto os demais, 12.326 foram usados para validação, na avaliação do desempenho do modelo.

Foi aplicada normalização a partir dos valores máximo e mínimo das amostras de treino para o intervalo $[-1, 1]$ e, por fim, treinada a rede perceptron multi-camada regressora, sem alteração dos parâmetros *default*. Dessa forma, foi utilizada função de ativação do tipo unidade linear retificada (ReLU), do inglês *rectified linear unit* e método de otimização *Adam* com base em gradiente estocástico, proposto por Kingma et al. (2017).

Os principais hiper-parâmetros utilizados foram:

Figura 11: Registro do procedimento de calibração do padrão desenvolvido.



Fonte: ISI-SIM (2022)

- Número de camadas ocultas: 100;
- Taxa de aprendizado (do inglês, *learning rate*): 10^{-3} .

Após treinado, o modelo pode ser utilizado para reprojeter os pontos de referência de validação extraídos da imagem de volta para o domínio real e, portanto, comparados com os valores conhecidos. Dessa forma, a métrica de validação foi este erro de reprojeção, considerando-se distância euclidiana.

4 RESULTADOS E DISCUSSÕES

Os resultados obtidos serão apresentados e discutidos neste capítulo. O método proposto foi avaliado em termos do reconhecimento do padrão desenvolvido e erro de reprojeção alcançado.

4.1 Ensaio estático

No ensaio estático exploratório, foi identificado que as transições de branco para preto, ao longo da linha de captura, variaram de 5 a 7 pixels, conforme ilustrado na Figura 12, um recorte da região central da imagem, onde as faixas de transição que caracterizam uma linha possuem valores menores.

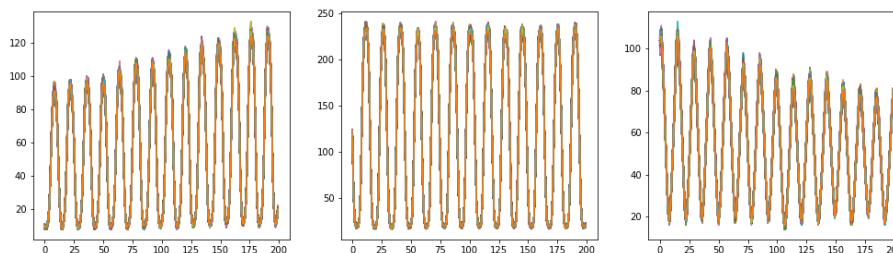
Figura 12: *Detalhe da região central de imagem capturado no ensaio estático.*



Fonte: autor (2023)

Além da diminuição da nitidez da imagem nas regiões mais distantes do centro do sensor, caracterizado pelo aumento do número de pixels de transição, pode-se observar a partir das imagens capturadas que a condição de iluminação deprecia nas laterais, próximas ao início e ao final do sensor da câmera. Conforme ilustrado na Figura 13, o contraste das regiões à esquerda e à direita são menores que na região central, onde há maior incidência luminosa.

Justifica-se, portanto, o projeto de tabuleiros distintos ao longo do campo de visão, de modo que a imagem de cada tabuleiro individualmente seja mais homogênea. Ainda assim, as linhas registradas na captura estão bem caracterizadas ao longo de toda a imagem, permitindo concluir que todo o campo de visão da câmera de linha pode ser utilizado para registrar características geométricas da ordem de milímetro.

Figura 13: *Curvas de cor, em escala de cinza, por linha de imagem do ensaio estático.*

Fonte: autor (2023)

4.2 Reconhecimento do padrão de calibração

A primeira etapa do algoritmo de calibração proposto consiste em extrair os dados de projeção dos pontos de referência na imagem. Portanto, foi avaliado o reconhecimento dos tabuleiros no conjunto de capturas realizadas. De dez imagens obtidas em condição ideal, ou seja, conforme descrito na Seção 3.2, apenas uma forneceu todo o conjunto de pontos, composto pelas posições internas das cinco *grids* utilizados, conforme ilustrado na Tabela 4.1.

Tabela 4.1: *Extração de pontos bem-sucedidas por imagem*

Captura	Tabuleiro 1	Tabuleiro 2	Tabuleiro 3	Tabuleiro 4	Tabuleiro 5
1	OK	OK	OK	OK	OK
2	OK	-	-	OK	OK
3	OK	-	-	-	OK
4	OK	-	-	OK	OK
5	OK	-	OK	OK	-
6	OK	OK	-	OK	OK
7	OK	OK	-	OK	OK
8	OK	OK	-	OK	OK
9	OK	OK	-	OK	OK
10	OK	OK	OK	OK	-

Os resultados evidenciam taxa de identificação satisfatória dos pontos de referência periféricos, de 100% e 80% nos tabuleiros 1 e 5, respectivamente. Também no 4, com 90%. Além disso, pelo menos uma captura retornou toda a informação necessária para aplicação do algoritmo de calibração. Dessa forma, a proposta de utilização de conjuntos distintos se mostrou eficaz, uma vez que seu objetivo era garantir a obtenção de dados das bordas da imagem, onde os efeitos de distorção são acentuados e o contraste era menor. Identificasse, neste caso, vantagem em relação à flexibilidade do método, ao permitir utilização dos tabuleiros de forma independente e, em caso de região danificada, não inutilizar o padrão.

Nos demais tabuleiros, 2 e 3, houve reconhecimento dos *grids* em 60% e 30%, respectivamente. O baixo desempenho observado, principalmente na região central da imagem, foi atribuído ao método de segmentação manual utilizado. Um ajuste fino dos parâmetros de recorte poderiam melhorar este resultado, uma vez que o recorte das imagens foi realizado com base em parâmetros fixos de posição 2D e, portanto, não há garantia de que o

tabuleiro de referência esteja centralizado corretamente no segmento obtido. É possível, também, que o elevado número de pontos nestes tabuleiros com maior quantidade de casas deixe o algoritmo de extração de cantos internos menos robusto.

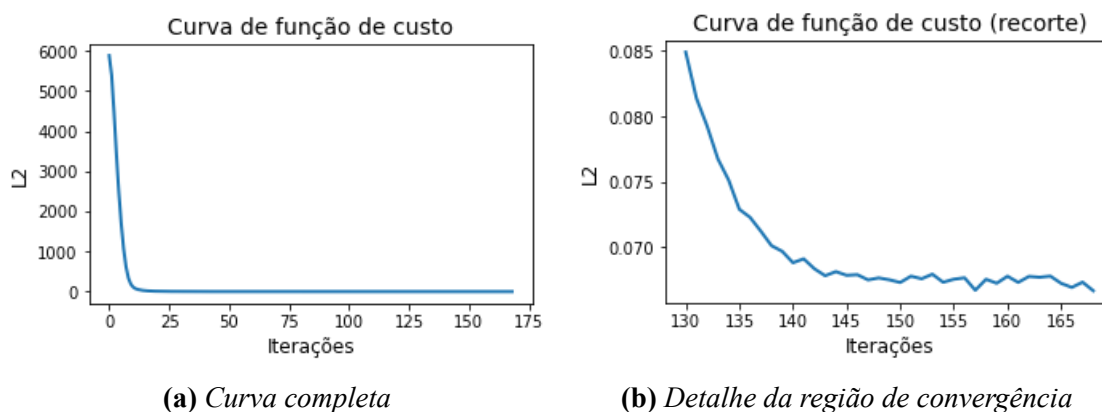
Apesar disso, a extração consistente dos pontos dos tabuleiros 1, 4 e 5, aliada ao conjunto de pontos completamente extraído da captura 1, foi suficiente para o avanço do trabalho em termos do algoritmo de calibração, uma vez que apenas uma imagem é requerida no método proposto. Ainda assim, serão enunciadas possibilidades de refinamento no Capítulo 5, em termos de trabalhos futuros.

De posse do conjunto total de coordenadas projetadas na captura 1, 41.085 pontos, foi desenvolvido o algoritmo de calibração com base em IA.

4.3 Regressor MLP

Na etapa de treinamento do modelo regressor MLP, foi analisado o comportamento da função de custo, do inglês *loss*. Ela representa a média do erro quadrático, do inglês *mean of squared errors* (MSE) e, portando, sua redução expressa melhoria do desempenho do regressor. Conforme ilustrado na Figura 14, o modelo foi capaz de regredir a partir dos dados fornecidos em 168 iterações, reduzindo a função de custo para 0.066655. Dado o número de amostras, 28.759 coordenadas de referência, e o tempo de processamento, 3 minutos com placa gráfica AMD Radeon RX 540, conclui-se baixo custo computacional do método.

Figura 14: Curva de função de custo durante o treinamento.



Fonte: autor (2023)

Para validação, foi realizada predição sobre conjunto de dados não utilizados durante o treinamento, *validation split*. O resultado obtido foi, em termos de coeficiente de determinação, denotado por R^2 , 0.999984, interpretado “como a proporção do desvio na variável dependente que é previsível a partir da variável independente”, de DataScience.eu (2023). É possível concluir que o treinamento foi bem sucedido, uma vez que o desempenho se manteve nos dados de validação.

A partir das informações estatísticas do erro, calculado pela distância euclidiana entre pontos de referência de teste, reais, e a reprojeção das coordenadas correspondentes extraídas da imagem, é possível quantificar a capacidade de medição dos sistema de visão.

Ao avaliar o erro de reprojeção alcançado, descrito na Tabela 4.2, foi observada dispersão significativa dos pontos. Calibrado, ele é capaz de fornecer posições com erro sistemático de aprox. 0.426 mm e precisão de aprox. 0.312 mm, totalizando um erro de medição, nos termos do Vocabulário Internacional de Metrologia (VIM), de aprox. 0.737 mm.

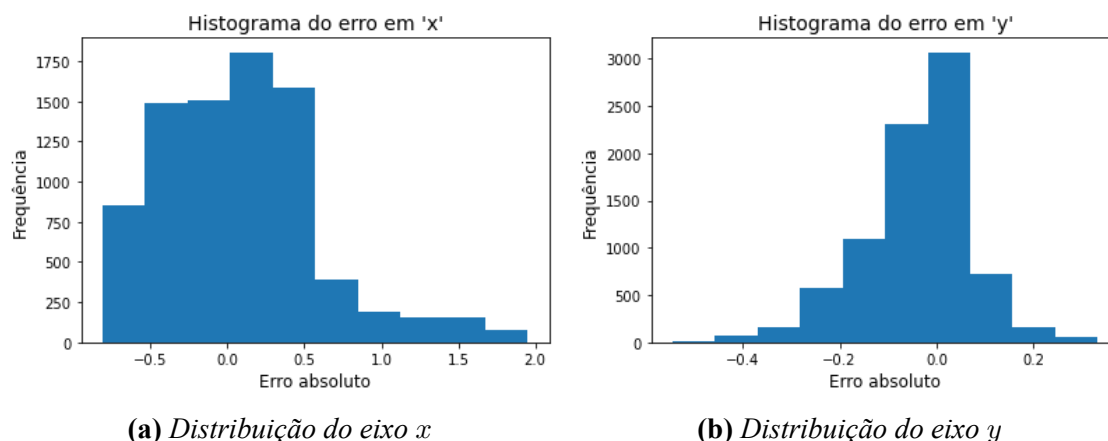
Tabela 4.2: Resultados estatísticos dos erros de reprojeção do conjunto de validação

Métrica	Média	Desvio padrão	Mediana	Percentil 95%	Máximo
Erro de reprojeção	0.425657	0.311758	0.366756	1.096964	1.951899

Uma análise dos eixos ortogonais de forma independente permite melhorar esse resultado, pois permite desconsiderar a parcela sistemática dos erros. Uma vez que os fenômenos relacionados à construção da imagem 2D diferem entre os eixos u e v , devido à utilização de câmera de linha, cabe avaliação independente dos erros de reprojeção para dimensões reais. Dessa forma, pode-se discutir fontes de contribuição de incerteza mais detalhadamente.

A Figura 15 ilustra a distribuição dos erros absolutos na reprojeção das coordenadas x e y separadamente. É notável a diferença entre os resultados obtidos em cada eixo, dado que o desvio padrão do erro em x é de 0.506604 mm e, em y , 0.113645 mm. Esses valores representam o erro aleatório de cada componente. Ao considerar as dimensões do padrão de calibração, de aprox. 100x460 mm, é possível descrever as contribuições de erro aleatórios em termos relativos como similares: em x , 0.1126%; enquanto em y , 0.1136%.

Figura 15: Histogramas de erro absoluto por eixo.



Fonte: autor (2023)

A principal contribuição de erro no eixo x , paralela ao sensor da câmera de linha, é causada pela deformação da lente devido à utilização de todo o campo de visão da câmera. Dessa forma, resta concluir que o modelo não foi capaz de descrever precisamente este fenômeno, ainda que selecionado pela aplicabilidade em sistemas com aspectos não-lineares. Considerando que o efeito de distorção da imagem é acentuado nas extremidades da captura, seu impacto é desprezível na maior parte da imagem. Isto significa que um número pequeno de amostras oferece informações de distorção e, portanto, nosso conjunto

de dados é desbalanceado nesse sentido. O impacto no processo de treinamento é que o modelo deu pouca prioridade para esse aspecto e, conseqüentemente, não o absorveu durante o aprendizado.

No eixo y , por sua vez, foram obtidos erros absolutos menores. Portanto, a utilização de um *encoder* no sistema de visão para sincronizar do disparo das imagens com a velocidade de movimento da esteira, ou elemento móvel se justificou. Tornou o sistema mais robusto, independente de variações na velocidade da esteira.

Demais contribuições têm impacto em ambas as coordenadas e, dessa forma, não justificam a diferença de resultado obtida. Pode-se mencionar: defeitos de fabricação do padrão de calibração, no sentido de não ser um tabuleiro perfeito, em termos de alinhamento e tamanho das casas; e desalinhamento entre a câmera e a plataforma de deslocamento linear, com efeito semelhante à distorção tangencial.

Destaca-se que analisando os eixos ortogonais da imagem em separado, é possível compensar a componente sistemática do erro em cada coordenada. Desconsiderado-os, portanto, a incerteza combinada dos eixos ortogonais pode ser expressa pela raiz da soma quadrática, resultando em 0.519194 mm. Para que fosse possível satisfazer o requisito de tolerância dimensional de 1 mm, o erro de medição do sistema não poderia ultrapassar 0.5 mm. Ainda que a regressão tenha apresentado resultado satisfatório em termos de convergência, sem apresentar *overfit*, o sistema não é capaz de realizar inspeção dimensional com tolerância de 1 mm. Isto porque o erro de medição obtido extrapola os limites definidos pela tolerância desejada. Portanto, a aplicação do método de calibração proposto se mostrou insatisfatório para a aplicação específica.

5 CONSIDERAÇÕES FINAIS

O trabalho consistiu na proposta e avaliação de método de calibração para sistema de visão composto por câmera de linha. Incorporando aspectos não-lineares do sistema, especialmente distorções da lente, a fim de permitir utilização de todo o campo de visão da câmera em aplicações de inspeção dimensional com tolerância de 1 mm.

Ajustes finos dos parâmetros do sistema de visão foram realizados a partir de ensaios exploratórios com os equipamentos disponíveis, emprestados do ISI-SIM. A partir do estudo de ferramentas de ML e suas aplicações, no estado da arte, foi selecionada um rede regressora perceptron multi-camadas (MLPRegressor) para solução do problema.

A montagem e configuração dos equipamentos do sistema de visão foi bem sucedida, uma vez que as imagens foram obtidas de acordo com o esperado, com resolução pixelar similar nos eixos u e v , sem efeito de *flickering*. Ainda que operacional, a descrição destes procedimentos contribui para a compreensão de sistemas de visão com câmera de linha e detalha uma de suas aplicações.

O trabalho demonstrou viabilidade técnica do padrão de calibração proposto, ao reportar identificação da totalidade dos pontos de referência projetados na imagem em pelo menos um dos ensaios. Dessa forma, fornece subsídio para explorar técnicas de calibração que contemplem distorções da lente, fenômeno acentuado nas extremidades da captura.

A regressão a partir dos dados obtidos foi bem-sucedida, uma vez que convergiu para solução com alto R^2 , de 0.999984, e apresentou resultados de validação similares aos de treinamento. Entretanto se mostrou ineficaz para aplicação de dimensionamento com tolerância de 1 mm, em função da variância dos erros de reprojeção em x , majoritariamente. Dessa forma, concluiu-se que o modelo de ML proposto não foi capaz de incorporar as distorções ópticas da lente, presentes principalmente neste eixo.

Apesar disso, os resultados permitem a continuação do trabalho por meio de alterações ou refinamento do algoritmo: *features* adicionais podem ser utilizados no treinamento, como padrões de borda e gradientes de direção; testes com outro modelo de ML; ou, ainda, aplicar estratégia de pesos nos dados de entrada, a fim de priorizar o aprendizado dos efeitos não-lineares da lente, compensando assim o desbalanceamento do *dataset*.

Além disso, uma comparação com resultados de uma abordagem analítica seriam interessantes, para avaliar quantitativamente a eficácia do método de calibração proposto. Os dados obtidos permitem extensão do estudo da aplicação de IA, justificado pelo ganho de praticidade, flexibilidade e robustez ao processo de calibração de sistema de visão com uso de algoritmo de ML.

REFERÊNCIAS

- BRADSKI, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- BURKOV, Andriy. *Machine learning engineering*. [S.l.]: True Positive Incorporated Montreal, QC, Canada, 2020. v. 1.
- CHAI, Junyi et al. Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, v. 6, p. 100134, 2021. ISSN 2666-8270. DOI: <<https://doi.org/10.1016/j.mlwa.2021.100134>>. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S2666827021000670>>>.
- DATASCIENCE.EU. *Coeficiente de Determinação*. 2023. Disponível em: <<<https://data-science.eu/pt/matematica-e-estatistica/coeficiente-de-determinacao/>>>.
- DILEO, G. et al. Uncertainty of line camera image based measurements. In: 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC). [S.l.: s.n.], mai. 2017. P. 1–6. DOI: <10.1109/I2MTC.2017.7969698>.
- DONNÉ, Simon et al. Robust plane-based calibration for linear cameras. In: 2017 IEEE International Conference on Image Processing (ICIP). [S.l.: s.n.], set. 2017. P. 36–40. DOI: <10.1109/ICIP.2017.8296238>.
- DUDA, Alexander; FRESE, Udo. Accurate Detection and Localization of Checkerboard Corners for Calibration. In: BRITISH Machine Vision Conference. [S.l.: s.n.], set. 2018. Disponível em: <<<https://api.semanticscholar.org/CorpusID:52286128>>>.
- GONZALEZ, Rafael C; WOODS, Richard E. *Processamento de imagens digitais*. 3. ed. [S.l.]: Pearson Education, 2010.
- GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. *Deep learning*. [S.l.]: MIT press, 2016.
- GUPTA, R.; HARTLEY, R.I. Linear pushbroom cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 19, n. 9, p. 963–975, set. 1997. ISSN 1939-3539. DOI: <10.1109/34.615446>.
- HARRIS, Charles R. et al. Array programming with NumPy. *Nature*, v. 585, p. 357–362, 2020. DOI: <10.1038/s41586-020-2649-2>.

- HEIKKILA, J. Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, n. 10, p. 1066–1077, out. 2000. ISSN 1939-3539. DOI: <10.1109/34.879788>.
- HEIKKILA, J.; SILVEN, O. A four-step camera calibration procedure with implicit image correction. In: PROCEEDINGS of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S.l.: s.n.], jun. 1997. P. 1106–1112. DOI: <10.1109/CVPR.1997.609468>.
- HORAUD, R.; MOHR, R.; LORECKI, B. On single-scanline camera calibration. *IEEE Transactions on Robotics and Automation*, v. 9, n. 1, p. 71–75, fev. 1993. ISSN 2374-958X. DOI: <10.1109/70.210796>.
- HUNTER, John D. Matplotlib: A 2D graphics environment. *Computing in science & engineering*, IEEE, v. 9, n. 3, p. 90–95, 2007.
- JABŁOŃSKI, Mirosław; TADEUSIEWICZ, Ryszard. Vision-based detection of events using line-scan camera. In: 2016 Second International Conference on Event-based Control, Communication, and Signal Processing (EBCCSP). [S.l.: s.n.], jun. 2016. P. 1–3. DOI: <10.1109/EBCCSP.2016.7605275>.
- KINGMA, Diederik P.; BA, Jimmy. *Adam: A Method for Stochastic Optimization*. [S.l.: s.n.], 2017. arXiv: <1412.6980> [cs.LG].
- LEE, Se-Na et al. Visual measurement of pile penetration and rebound movement using a high-speed line-scan camera. In: PROCEEDINGS 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292). [S.l.: s.n.], mai. 2002. v. 4, 4307–4312 vol.4. DOI: <10.1109/ROBOT.2002.1014436>.
- LIAO, Ruiying et al. Flexible calibration method for line-scan cameras using a stereo target with hollow stripes. *Optics and Lasers in Engineering*, v. 113, p. 6–13, 2019. ISSN 0143-8166. DOI: <<https://doi.org/10.1016/j.optlaseng.2018.09.014>>. Disponível em: <<<https://www.sciencedirect.com/science/article/pii/S0143816618308984>>>.
- LUNA, C.A. et al. Method to measure the rotation angles in vibrating systems. *IEEE Transactions on Instrumentation and Measurement*, v. 55, n. 1, p. 232–239, fev. 2006. ISSN 1557-9662. DOI: <10.1109/TIM.2005.860868>.
- LUNA, Carlos A. et al. Calibration of Line-Scan Cameras. *IEEE Transactions on Instrumentation and Measurement*, v. 59, n. 8, p. 2185–2190, ago. 2010. ISSN 1557-9662. DOI: <10.1109/TIM.2009.2031344>.
- LUTZ, W. et al. Fast measuring of solid angles by means of CCD line scan cameras with analytically rectified optics. In: PROCEEDINGS of 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. [S.l.: s.n.], nov. 1994. v. 2, 1007–1008 vol.2. DOI: <10.1109/IEMBS.1994.415296>.
- MCKINNEY, Wes et al. Data structures for statistical computing in python. In: AUSTIN, TX. PROCEEDINGS of the 9th Python in Science Conference. [S.l.: s.n.], 2010. v. 445, p. 51–56.

- NIU, Menghui et al. The Line Scan Camera Calibration Based on Space Rings Group. *IEEE Access*, v. 6, p. 23711–23721, 2018. ISSN 2169-3536. DOI: <10.1109/ACCESS.2018.2817629>.
- PEDREGOSA, Fabian et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research*, v. 12, Oct, p. 2825–2830, 2011.
- PETTY, R S; ROBINSON, M; EVANS, J P O. 3D measurement using rotating line-scan sensors. *Measurement Science and Technology*, v. 9, n. 3, p. 339, mar. 1998. DOI: <10.1088/0957-0233/9/3/005>. Disponível em: <<<https://dx.doi.org/10.1088/0957-0233/9/3/005>>>.
- ROSA, Artur Martini da. *Simulação computacional de um dispositivo robótico com sistema de visão baseado em Deep Learning*. 2021. Trabalho de conclusão de graduação – UFRGS.
- ROSENBLATT, Frank. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, American Psychological Association, v. 65, n. 6, p. 386, 1958.
- RUSSEL, Stuart; NORVIG, Peter. *Artificial intelligence: a modern approach.*–1136 p. [S.l.]: New Jersey: Prentice Hall, 2020.
- STEGER, Carsten; ULRICH, Markus. A Camera Model for Line-Scan Cameras with Telecentric Lenses. *International Journal of Computer Vision*, v. 129, jan. 2021. DOI: <10.1007/s11263-020-01358-3>.
- STEGER, Carsten; ULRICH, Markus; WIEDEMANN, Christian. Machine vision algorithms and applications. In: [s.l.]: John Wiley & Sons, 2018. 3.9.3.4 Nonlinearities of the Line Scan Camera Model.
- VOULODIMOS, Athanasios et al. Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, Hindawi, v. 2018, 2018.
- WANG, Qing et al. Computer vision for yarn microtension measurement. *Applied Optics*, v. 55, p. 2393, mar. 2016. DOI: <10.1364/AO.55.002393>.
- WENDEL, Alexander; UNDERWOOD, James. Extrinsic Parameter Calibration for Line Scanning Cameras on Ground Vehicles with Navigation Systems Using a Calibration Pattern. *Sensors*, v. 17, n. 11, 2017. ISSN 1424-8220. DOI: <10.3390/s17112491>. Disponível em: <<<https://www.mdpi.com/1424-8220/17/11/2491>>>.
- YING, Xianghua et al. A direct method for the extrinsic calibration of a camera and a line scan LIDAR. In: 2014 IEEE International Conference on Mechatronics and Automation. [S.l.: s.n.], ago. 2014. P. 571–576. DOI: <10.1109/ICMA.2014.6885760>.
- ZHANG, Jiabin et al. A New Calibration Method of Line Scan Camera for High-Precision Two-Dimensional Measurement. In: 2018 IEEE 14th International Conference on Automation Science and Engineering (CASE). [S.l.: s.n.], ago. 2018. P. 678–683. DOI: <10.1109/COASE.2018.8560438>.
- ZISSERMAN, Richard Hartley Andrew. *Multiple view geometry in computer vision*, 2004.