

**UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE CIÊNCIAS EXATAS E TECNOLOGIA
CURSO DE SISTEMAS DE INFORMAÇÃO**

PAULO VÍTOR LIBÓRIO PEREIRA

**ABORDAGEM YOLOv5 PARA DETECÇÃO E CLASSIFICAÇÃO DE
ESFERAS DE SOLDA NO ENCAPSULAMENTO DE
SEMICONDUCTORES**

Itacoatiara – Amazonas

Fevereiro – 2023

PAULO VÍTOR LIBÓRIO PEREIRA

**ABORDAGEM YOLOv5 PARA DETECÇÃO E CLASSIFICAÇÃO DE
ESFERAS DE SOLDA NO ENCAPSULAMENTO DE
SEMICONDUCTORES**

Monografia apresentada ao Instituto de Ciências Exatas e Tecnologia da Universidade Federal do Amazonas como parte dos requisitos necessários para a obtenção do título de Bacharel em Sistemas de Informação.

FELIPE GOMES DE OLIVEIRA

Itacoatiara – Amazonas

Fevereiro – 2023

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

P436a Pereira, Paulo Vítor Libório
Abordagem YOLOv5 para detecção e classificação de esferas de solda no encapsulamento de semicondutores / Paulo Vítor Libório Pereira . 2023
18 f.: il. color; 31 cm.

Orientador: Felipe Gomes de Oliveira
TCC de Graduação (Sistemas de Informação) - Universidade Federal do Amazonas.

1. YOLOv5. 2. Detecção de objetos. 3. Classificação de imagens. 4. Semicondutores. 5. Aprendizado profundo. I. Oliveira, Felipe Gomes de. II. Universidade Federal do Amazonas III. Título

Abordagem YOLOv5 para Detecção e Classificação de Esferas de Solda no Encapsulamento de Semicondutores

Paulo Vítor Libório Pereira, Felipe Gomes de Oliveira

Instituto de Ciências Exatas e Tecnologia – Universidade Federal do Amazonas
(ICET/UFAM) – Itacoatiara – Amazonas – Brasil

vitorliborio3@gmail.com, felipeoliveira@ufam.edu.br

Resumo. A detecção de objetos baseada em visão computacional é essencial para acelerar a produção de produtos eletrônicos. Entretanto, a detecção automática de defeitos em superfícies de PCB ainda é uma tarefa desafiadora. Apesar da existência de diversos detectores baseados em visão computacional que abordam esses problemas, os detectores atuais enfrentam desafios para atingir alta precisão de detecção e velocidade. Para o treinamento e teste da rede neural, foram consideradas três métricas para avaliar os resultados da detecção: precisão, recall e mAP e para a classificação considerou-se a acurácia média. O objetivo consiste em propor uma abordagem para detectar e classificar três categorias de esferas de solda, no processo de soldagem de pastilhas de silício em BGA contidos em substratos PCB combinando o modelo YOLOv5 e uma CNN. Os resultados experimentais mostram que o detector alcançou um desempenho considerável, marcando um mAP@50 de 92,6% para o modelo YOLOv5 e uma acurácia média de 97,87% para o modelo de CNN.

1. Introdução

A produção de dispositivos de memória representa um grande desafio de fabricação em relação à dimensão dos componentes ser de tamanho ínfimo, além de exigir alta precisão em sua operação. Nesse contexto, a inspeção de semicondutores tem a função de informar sobre os erros específicos de fabricação, que podem estar relacionados com problemas de produção. Assim sendo, a inspeção de semicondutores inclui a avaliação da ocorrência de falhas relacionadas aos materiais envolvidos nos processos ou definições inadequadas de parâmetros (Zhang et al., 2021).

Relatos recentes mostram que as inspeções de qualidade continuam sendo realizadas em sua maioria por humanos. Apesar de profissionais qualificados, essas análises apresentam eficácia entre 80% e 90% dos casos (Silva et al., 2022). No entanto, esse processo manual é repetitivo e trabalhoso e sujeito a variações e fadigas. Na Figura 1, é apresentado um exemplo de inspeção visual humana de pastilhas de silício no qual o operador deve percorrer visualmente todos os componentes do substrato da Placa de Circuito Impresso (PCB) em busca de diversos tipos de defeitos, sendo, portanto, um processo desafiador em relação ao tamanho das esferas de solda e as características envolvidas na análise do problema.

Com os avanços da inteligência artificial, os algoritmos de aprendizagem profunda têm melhorado os resultados em diversas aplicações (Khan et al., 2020). Os referidos algoritmos podem ser usados para obter uma inspeção mais confiável e eficiente em cenários altamente complexos. Em comparação aos métodos tradicionais, esses benefícios do aprendizado profundo motivam os pesquisadores a desenvolver algoritmos para analisar as imagens capturadas das fábricas e fornecer metodologias de inspeções inteligentes. Para estefim, a detecção de objetos tem atraído muita atenção nos últimos anos (Zhao et al., 2019).

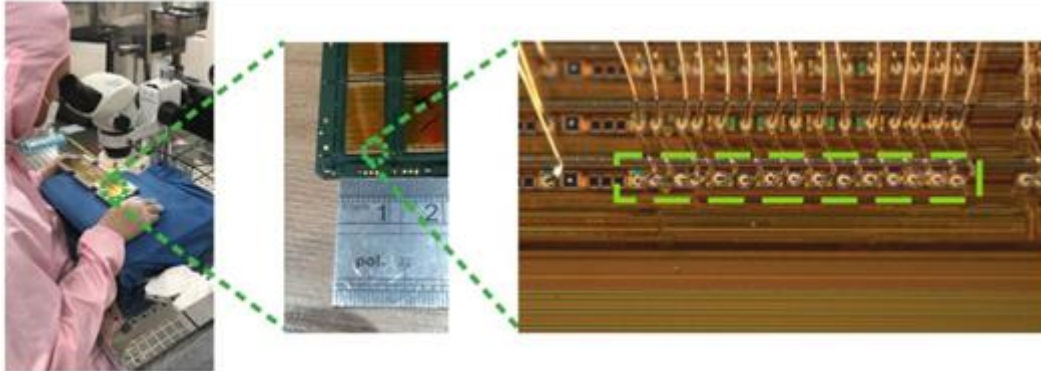


Figura 1. Inspeção visual convencional no processo de soldagem de pastilhas de silício

A detecção de objetos visa detectar um objeto em uma imagem, localizá-lo com uma caixa delimitadora e classificá-lo com a pontuação de confiança para a classe prevista. É um dos problemas fundamentais em aplicações de visão computacional e fornece informações valiosas (Cao et al., 2017). Além disso, a detecção de objetos pode ajudar no processo de inspeção de fabricação, detectando o objeto desejado nos quadros de vídeo e decidindo se a inspeção foi realizada corretamente ou não.

Em relação à metodologia, este trabalho foi desenvolvido em quatro etapas. A primeira etapa consistiu na aquisição de imagens das pastilhas de silício. A segunda etapa consistiu no pré-processamento das imagens, onde foram aplicadas técnicas de aumento de dados e rotulação das caixas delimitadoras. A terceira etapa compreendeu o processo de treinamento do conjunto de dados para a detecção das caixas delimitadoras. A quarta etapa integra o processo de classificação das categorias de solda por meio da detecção e localização das caixas delimitadoras nas imagens.

Este estudo se propõe a colaborar na resolução do problema de inspeção automática no controle de qualidade do processo de soldagem de pastilhas de silício, chamados die, em BGA (do inglês, *Ball Grid Array*). Para isso, é apresentada uma abordagem para detectar e classificar esferas de solda em três categorias: i) Correta; ii) Falha; e iii) Ausente. Nesse processo foram utilizadas técnicas de visão computacional: i) na detecção de regiões onde deveriam ser encontradas esferas de solda, usando uma Rede Convolutiva da família YOLO (do inglês, *You Only Look Once*); e ii) na classificação de condições de solda, usando um modelo de Rede Neural Convolutiva (CNN).

Os resultados obtidos mostram que o algoritmo pode identificar as categorias que representam todos os tipos de condições de solda abordadas. A precisão da etapa de detecção atinge um $mAP@50$ de 92,60% e acurácia média de 97,87% para a classificação utilizando um modelo de CNN. Portanto, os resultados indicam que o YOLOv5 combinado com a CNN pode ser uma alternativa para apoiar especialistas no processo de inspeção industrial.

O restante do artigo está organizado da seguinte maneira. A Seção 2 apresenta alguns conceitos básicos e discute os trabalhos relacionados. A Seção 3 apresenta o método

de pesquisa utilizado, enquanto a Seção 4 mostra os resultados e as discussões. A Seção 5 apresenta as conclusões e os trabalhos futuros.

2. Fundamentação Teórica

2.1. Conceitos Relacionados

Esta seção apresenta os principais conceitos teóricos utilizados no trabalho proposto.

2.1.1 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNNs), descritas por (LeCun et al., 2001), são uma arquitetura comum de Aprendizado Profundo utilizada na resolução de problemas de Visão Computacional, especialmente para detecção de objetos. Em resumo, CNNs são redes neurais que aplicam, pelo menos em uma camada, uma operação chamada Convolução (Googfellow et al., 2016).

Dessa forma, uma CNN é formada por diversas camadas, cada uma com uma função específica no processo de aprendizado e classificação de imagens. As principais camadas são: Camada de Convolução e Camada de *Pooling*, responsáveis pela abstração e aprendizado da rede; e Camada Totalmente Conectada, que é responsável por gerar a classificação das imagens com base no aprendizado obtido durante o treinamento. A função de ativação ReLU é geralmente inserida entre as camadas de Convolução e *Pooling*, e ajuda no processo de treinamento do modelo (Yamashita et al., 2018).

2.1.2 Detecção de Objetos com YOLOv5

Nos últimos anos, a detecção de objetos com base em aprendizado profundo tem sido uma área importante de pesquisa. Para melhorar a identificação de imagens por meio de métodos de inteligência artificial, é necessário não apenas se concentrar no processo de classificação de diferentes tipos de imagens, mas também estimar com precisão quais objetos estão presentes em cada figura a ser processada e qual é a sua localização. Essa tarefa é conhecida como detecção de objetos, conforme mencionado por (Zhao et al., 2019). Os algoritmos detectores são definidos por (Malta et al., 2021) como redes neurais profundas, especificamente redes neurais convolucionais (CNN).

Dentre os diversos algoritmos existentes para detecção e reconhecimento de imagens, as séries R-CNN e YOLO são as mais destacadas. Embora a série R-CNN tenha melhor acurácia, sua velocidade de detecção é muito mais baixa do que a apresentada pela YOLOv5, o que a torna inapropriada para tarefas que exigem rapidez, como o monitoramento em tempo real de registros de vídeo, conforme afirmado por (Jia Yao s.d.). O trabalho de (Jin e Niu 2021) explica que a série de algoritmos YOLO utiliza a ideia de regressão para aprender as características visuais gerais de um alvo de maneira simples, otimizando o tempo necessário para resolver o problema. A YOLO utiliza redes neurais de estágio único para detectar completamente o posicionamento e a classificação dos objetos em análise.

Segundo (Zhao et al., 2019), a principal ideia da YOLO é usar a imagem inteira como entrada da rede e retornar diretamente o posicionamento de uma *bounding box* e a categoria à qual a *bounding box* pertence como possíveis saídas.

2.1.3 Classificação de Objetos

A classificação de uma imagem é uma tarefa desempenhada por um algoritmo com a finalidade de, dada uma série de classes, definir qual delas melhor se encaixa aos padrões analisados. A localização, por sua vez, torna o algoritmo responsável por criar uma caixa delimitadora na posição que o objeto se encontra na imagem. Contudo, o problema envolvendo a detecção abrange múltiplos objetos, envolvendo diversas classes em uma mesma imagem, no qual o algoritmo deve ser capaz de classificá-los e localizá-los de modo simultâneo (Redmon et al., 2016).

2.1.3 Técnicas de Divisão

A técnica estatística conhecida como Validação Cruzada é usada para validar modelos a partir de um conjunto de dados. Essa técnica envolve dividir os dados em dois grupos aleatórios: um para teste e outro para treinamento. O conjunto de treinamento é então dividido em dois subgrupos: um é usado para escolher o modelo mais adequado e o outro é usado para validar o modelo selecionado.

O objetivo dessa divisão é testar um novo paradigma, utilizando um conjunto de dados diferente daquele utilizado para estimar os parâmetros. Essa abordagem permite a escolha entre vários candidatos para um novo paradigma. No entanto, essa escolha pode resultar em um ajuste excessivo do conjunto de validação, prejudicando a generalização. Para garantir uma boa generalização, o novo paradigma selecionado deve ser avaliado em um conjunto de teste diferente do subconjunto de validação. Esse método ajuda a evitar o ajuste excessivo e permite uma validação mais confiável dos modelos. (Haykin, 2001).

A técnica de validação cruzada K-fold é frequentemente utilizada quando há um número limitado de dados classificados. Nessa situação, a validação cruzada múltipla é uma opção viável, que envolve dividir um conjunto de N objetos em K subconjuntos, onde K é maior que 1 e é divisível por N.

No método K-fold, o novo paradigma é treinado com todos os subconjuntos, exceto um, que é utilizado para verificar o erro de validação. Esse processo é repetido K vezes, com um subconjunto diferente sendo usado para validação em cada iteração. A medida de desempenho desse método é a média do erro quadrado encontrado na validação de todas as tentativas do experimento.

2.2. Trabalhos Relacionados

O problema considerado no escopo deste trabalho já foi abordado na literatura com diferentes técnicas.

Em (Silva et al., 2022) foi proposta uma rede neural convolucional composta por duas camadas convolucionais, com 16 filtros na primeira camada e 32 filtros na segunda camada. A função de ativação ReLU é usada em camadas convolucionais com *Max Pooling*. Na etapa de treinamento, o algoritmo de otimização SGD foi aplicado, com taxa de aprendizado igual a 0,001, por 40 épocas e usando um tamanho de lote igual a 32. A CNN foi utilizada visando a contribuição no problema de controle de qualidade no procedimento

de soldagem de pastilhas de silício, permitindo a detecção de erros antes do processo de moldagem de resina epóxi. Dessa forma, gerando informações úteis para corrigir configurações de equipamentos e prever falhas das matérias-primas e insumos utilizados no processo. A metodologia proposta é composta de duas etapas: i) segmentação da esfera de solda utilizando o método *Haar Cascade*; e ii) classificação de esferas de solda por meio de aprendizado profundo. Os resultados mostraram a precisão obtida de 99,4% usando um modelo de classificação de rede neural convolucional. Além disso, a abordagem proposta apresentou alta acurácia, de 92,8% e 75,7%, mesmo na presença dos ruídos aditivos *Salt and Pepper* e Gaussiano, nos piores cenários respectivamente.

A inspeção de placas de circuito impresso (PCB) tem sido uma das tarefas importantes na indústria eletrônica. Em (Klco et al., 2022) o autor trata da inspeção visual automatizada de placas eletrônicas na fabricação seriada de dispositivos eletrônicos fornecida usando a rede neural YOLO. A rede neural YOLO foi treinada em um conjunto de dados próprio, segundo os autores o método implementado levou a uma maior confiabilidade no processo de fabricação. O método YOLOv5 foi escolhido devido a velocidade de detecção e precisão. Os resultados mostraram precisão de 92,7%, recall de 87,9% e pontuação média de precisão mAP de 90,6%.

Visando criar uma ferramenta de inspeção eficiente (Zhang et al., 2021) propuseram um sistema de inspeção de juntas de solda baseado em visão de máquina, que pode detectar o status das juntas de solda e fornecer os resultados atuais da soldagem aos trabalhadores. No pré-processamento de imagem, é utilizado o algoritmo LSD para detectar uma linha reta em uma área específica. A imagem é pré-processada por rotação, o algoritmo LSD e a transformada de *Hough* são usados para localizar as juntas de solda, e os métodos baseados em área, como a proporção de vários pixels coloridos, são usados para detecção. Um total de 100 conjuntos de bombas de água foram testados nesse experimento, e produtos com diferentes status de inspeção de juntas de solda aparecem aleatoriamente. Sob aproximadamente o mesmo ambiente de iluminação na oficina, a taxa de precisão final do experimento foi de 98,00%, o que verificou a viabilidade e praticabilidade do projeto.

No estudo de (Zhou et al., 2021) foi apresentado, um sistema de reconhecimento gráfico que pode quantificar a imagem de digitalização do verso e assim distinguir exceções. O objetivo desse trabalho foi monitorar o verso de pastilhas de silício. O método de monitoramento inclui inspeção visual, inspeção macroscópica, inspeção microscópica e inspeção hiperfina. Todos esses métodos, cruzados por diferentes características do processo, constroem um sistema de monitoramento de vários estágios. Como resultado, a verificação pode constatar e eliminar uma falha após a comparação com o *Backside Database* (BDB).

No trabalho de (Cao., 2021) é apresentado um sistema de inspeção visual, com o auxílio de robôs, combinada com o controle do programa de inspeção visual de forma que ambos trabalhassem juntos, para concluir uma tarefa de inspeção de processamento de alvo de metal do semicondutor de alta pureza. O objetivo era detectar os defeitos da superfície de alvos de metal em semicondutores. Os resultados demonstraram que o sistema pôde substituir completamente o teste manual tradicional, e assim, melhorar a qualidade de

usinagem e a eficiência do alvo de metal do semicondutor.

No contexto mais direcionado a inspeção de semicondutores, os autores (Srivastava et al., 2016) demonstram em seu trabalho, uma estratégia de inspeção não tradicional para monitoramento em linha aplicada em cenários de excursão usando a ferramenta 29xx. O principal objetivo, foi monitorar e rastrear um grande número de pastilhas afetadas por situações fora de controle. O método utilizou recursos 29xx de alto rendimento (TPT) combinado com uma metodologia para filtrar um grande volume de pastilhas para monitorar a variabilidade no processo e identificar pastilhas em risco. Os resultados demonstram, que essa metodologia permitiu a inspeção de aproximadamente 15 pastilhas por hora. Essa mesma metodologia foi subsequentemente implementada com sucesso para verificar/monitorar vários cenários de excursão em diferentes etapas do processo.

Os autores (D. Zhou, G. Zhang e Y. Guo, 2017) abordam em seu trabalho, um método de detecção de multi-interface da imagem utilizando-se da separação óleo-água com base no método de estatística de diferença. Teve como objetivo, propor um método baseado em estatísticas de diferença para detectar interfaces múltiplas *in vitro*. A metodologia é baseada no aprimoramento da interface da imagem óleo-água, posteriormente, é calculado o valor médio de cinza da linha da imagem óleo-água de acordo com o valor extremo local da diferença. Os resultados mostram que o método de detecção multi-interface tem alta precisão e atende aos requisitos de aplicações industriais, que podem ser usados para detecção multi-interface de cena semelhante.

Neste artigo, abordaremos a aplicação da arquitetura YOLOv5 e um modelo de CNN para detectar e classificar esferas de solda contidas em substratos de placas de circuito impresso (PCB), usando imagens coletadas do processo de soldagem de pastilhas de silício. YOLOv5 é um algoritmo de detecção de objetos, baseado em aprendizado profundo, que aprende padrões com base na profundidade da rede e no número de parâmetros aprendidos (Jocher et al., 2020).

3. Materiais e Métodos

O problema de inspeção da qualidade de solda de placas PCB foi abordado no escopo deste trabalho como uma tarefa de detecção e classificação mediante aprendizado supervisionado com a arquitetura YOLO e uma CNN. A metodologia adotada neste trabalho é dividida em quatro etapas principais: i) Aquisição de imagens; ii) Pré-Processamento; iii) Detecção; e iv) Classificação. Por fim, com vistas ao objetivo de avaliar e comparar a performance do algoritmo YOLOv5 para detecção de objetos algumas métricas de avaliação foram utilizadas. Cada etapa do desenvolvimento dessa pesquisa é resumida no fluxograma apresentado na Figura 2.

A Figura 2 apresenta a visão geral sobre o processo de detecção automática de esfera de solda no processo de soldagem de pastilhas de silício em *Ball Grid Array* (BGA), contidas em um substrato de placas de circuito impresso. Inicialmente, um conjunto de imagens de pastilhas de silício é captado por uma câmera. Nessas imagens, são apresentadas três categorias para a análise da classificação: correta, ausente ou falha. Após a construção do banco de imagens, todas as imagens foram divididas em um conjunto de treinamento, um

conjunto de validação e um conjunto de testes. Os conjuntos de treinamento e validação foram utilizados para o treinamento e o conjunto de testes foi usado para avaliação do modelo.

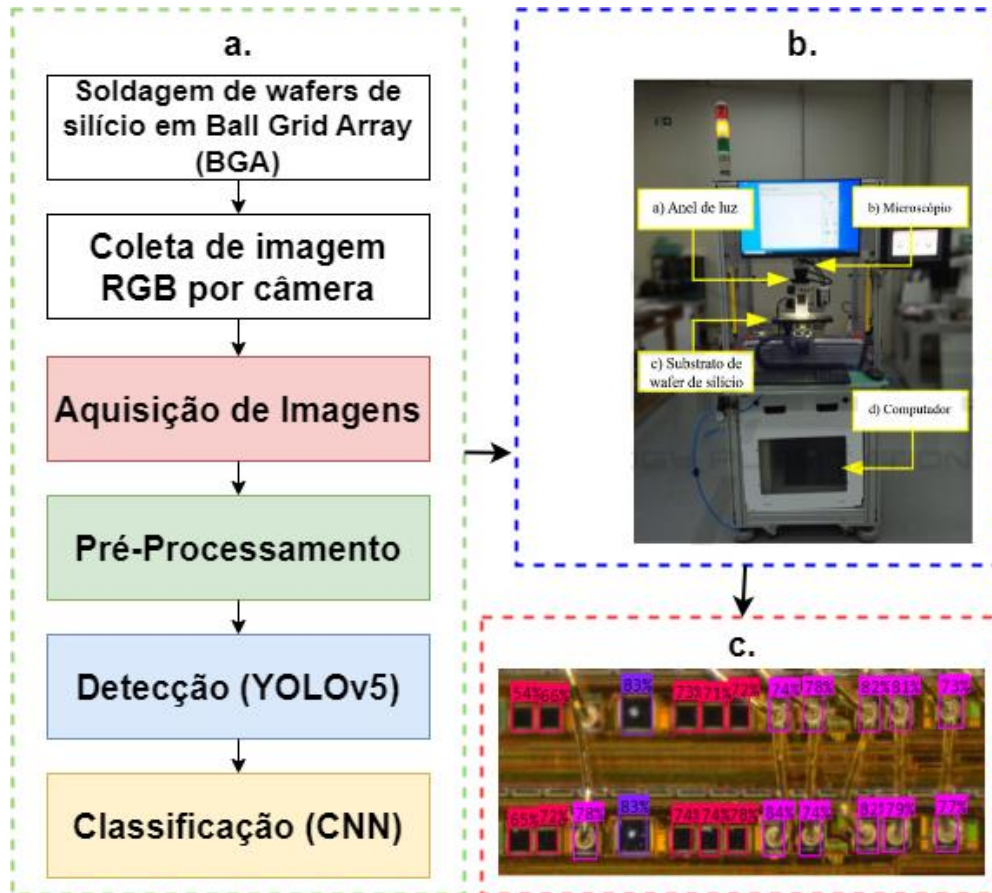


Figura 2. Metodologia proposta. a) Fluxograma do processo experimental. b) Plataforma de captura de imagens. c) Imagens do resultado do processo de detecção usando o YOLOv5

3.1 Aquisição de Imagem

Esse processo consiste na captura de um conjunto de imagens adquiridas no modelo *Red, Green, Blue* (RGB). Para isso, foi utilizado um dispositivo de aquisição de imagem, composto por um Estereomicroscópio Olympus SZ61TR, acoplado a uma câmera SC180. Essa estrutura dispõe de um sistema de iluminação que fornece luzes adequadas para a captura das imagens. Cada imagem capturada é encaminhada para o computador, de modo que seja processada posteriormente. A procedência das imagens originou-se do projeto de pesquisa e desenvolvimento do trabalho de (Silva et al., 2022) realizado em uma fábrica de componentes eletrônicos da Zona Franca de Manaus. O arcabouço experimental pode ser visualizado na Figura 2b.

3.2 Pré-Processamento de Imagem

A etapa de pré-processamento consiste em preparar as imagens das pastilhas de silício para realizar o treinamento do modelo YOLOv5. Para isso, as imagens foram redimensionadas para a resolução de 640x640 pixels, conforme as especificações da

arquitetura *YOLO* disponível no repositório (Jocher et al., 2020). Em seguida, foi empregada a técnica de aumento de dados (do inglês, *data augmentation*), utilizando a ferramenta *Roboflow* (Alexandrova et al. 2015), para suplementar os dados de treinamento aplicando as operações de girar (verticalmente e horizontalmente), giro de 90°, cortar (para descartar aleatoriamente um subconjunto da imagem original), rotação aleatória, tesoura (para distorcer aleatoriamente a imagem original em seu eixo vertical ou horizontal), escala de cinza (8 bits), matiz, saturação, brilho, exposição (para ficar mais claro ou escuro), desfoque (borramento) e ruído aleatório. As operações de aumento de dados são mostradas na Figura 3. Esse procedimento cria técnicas de aumento de dados para melhorar a diversidade do conjunto de treinamento e aprimorar a generalização do modelo (Sakkos et al., 2019). Cada imagem foi anotada utilizando a ferramenta *Roboflow*, construindo assim, uma coleção de referência com um total de 241 imagens para o conjunto de dados de treinamento e teste do modelo, conforme mostra a Figura 4, denominada SILICON_WAFERS_BGA-Dataset. A Tabela 1 apresenta o número total de categorias anotadas. Em seguida, as imagens foram exportadas no formato YOLOv5 PyTorch.

Tabela 1. Número total de categorias anotadas por classe.

Categorias	Quantidade
Correta	1444
Falha	646
Ausente	457
Total	2547

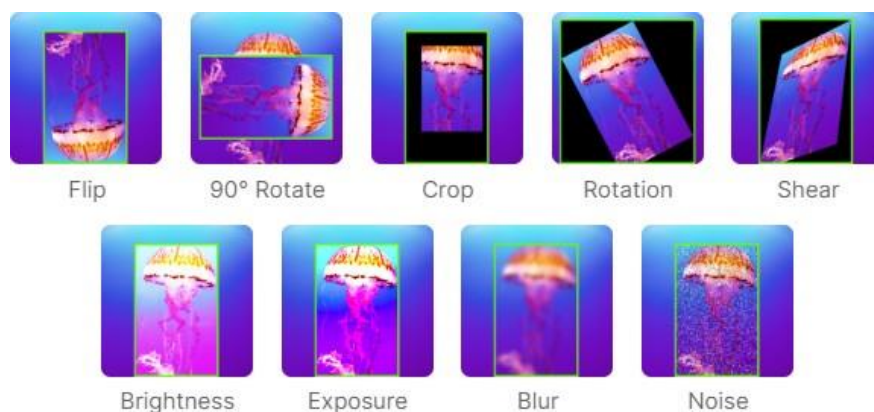


Figura 3. Operações de aumento de dados no Roboflow para suplementar os dados de treinamento

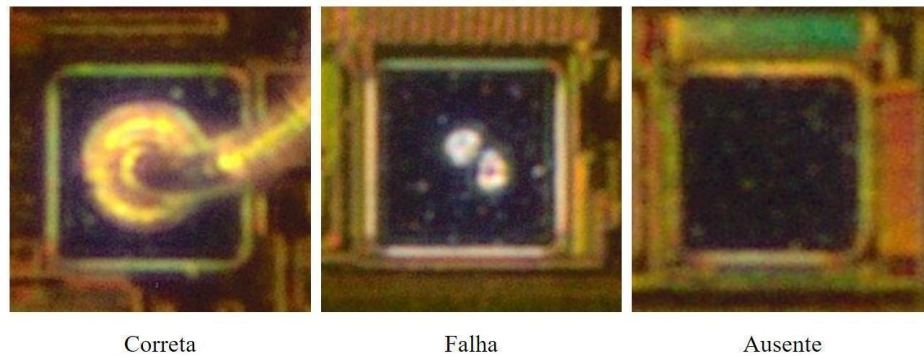


Figura 4. Amostra de imagens do conjunto de dados, dividido em Correta, Falha e Ausente

3.3 Detecção das Esferas de Solda (YOLOv5)

O *YOLOv5* é a quinta versão da série *You Only Look Once (YOLO)* de algoritmos para detecção de objetos (Jocher et al., 2020). Assim como seus predecessores, é uma rede de detecção de estágio único que a torna mais rápida em comparação com outros algoritmos de detecção de objetos (Yan et al., 2021), (Zhou et al., 2021). O modelo de reconhecimento do YOLOv5 pode garantir alta precisão na velocidade de detecção em tempo real.

De acordo com o tamanho da profundidade da rede e a largura do gráfico de recursos, a variação *Small (YOLOv5s)* é adotada como modelo de uso (Jocher et al. 2016), constituída por cinco modelos que diferem no tamanho da rede neural e que possuem integração nativa com o framework *PyTorch*. O modelo escolhido foi utilizado por dar ênfase nas três subdivisões principais da rede: i) *Backbone*: responsável por extrair características (*features*) da imagem; ii) *Neck*, onde são tratadas essas *features* para a etapa de detecção, agregando e combinando as mesmas; e iii) *Detect* (também chamado de *head*) que é a etapa de detecção, onde possui como saída um vetor contendo os dados das detecções feitas (descrições das caixas delimitadoras) resultando com as imagens e a localização do objeto de interesse formada. Na Figura 5, é apresentado o fluxo de dados entre essas três subdivisões.

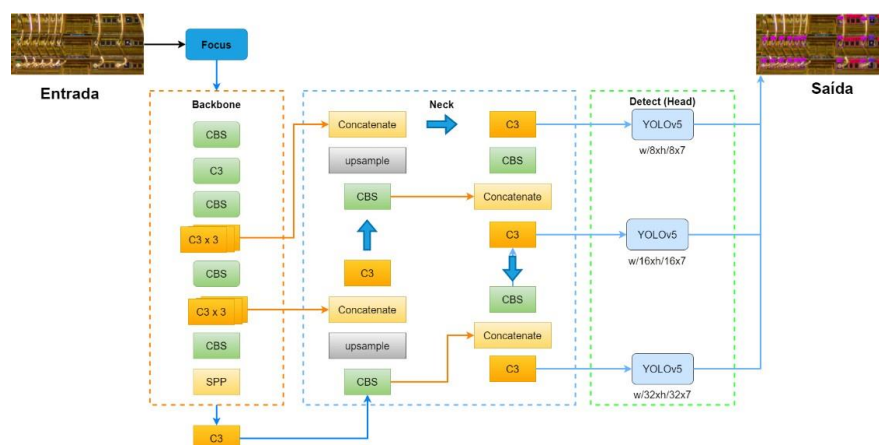


Figura 5. Arquitetura da rede YOLOv5s, destacando a função de cada subdivisão do modelo

Alguns hiperparâmetros significativos foram mantidos no estado padrão disponível no repositório, como o tamanho da entrada do modelo em três canais (RGB) de dimensões 640x640, o otimizador SGD, o limiar de *Non-maximum Supression* (NMS), responsável por evitar detecções múltiplas para um mesmo objeto (Redmon et al., 2016), com valor de 0,45 e o limiar de confiança da detecção em 0,25. A etapa de treinamento foi realizada na plataforma online *Google Colab* que disponibiliza uma máquina virtual com placa de vídeo dedicada (GPU), que nesse caso foi uma Tesla T4. Os hiperparâmetros foram selecionados por meio de observação empírica, utilizando uma abordagem de tentativa e erro, além de considerar os valores padrões utilizados pelo método YOLOv5 na base de imagens COCO. Os hiperparâmetros usados na versão de treinamento são mostrados na Tabela 2.

Tabela 2. Hiperparâmetros usados na versão de treinamento.

Hiperparâmetro	Descrição	Valor
Tamanho de imagem	Tamanho da imagem em pixels	640 x 640
Tamanho do <i>batch</i>	Números de amostras trabalhadas	16
Épocas	Número de vezes em que o algoritmo de treinamento foi executado em todo o conjunto de dados	300

3.4 Classificação das Esferas de Solda (CNN)

A partir das detecções, é realizado o processo de classificação para identificar as diferentes condições das esferas de solda: i) Correto; ii) Falha; e iii) Ausente. Nesta abordagem, é proposto um modelo de CNN para classificação de esferas de solda, baseado em aprendizado profundo. A CNN pode aprender representações eficientes para aplicação de classificação ou regressão em problemas reais. O modelo proposto tem duas camadas convolucionais com 16 filtros na primeira camada e 32 filtros na segunda camada. O tamanho dos filtros na primeira camada foi (5,5), enquanto na segunda camada foi (9,9). Este modelo usa ReLU como função de ativação, *Max Pooling* de janela de tamanho (2,2), e foi treinado com SGD, com taxa de aprendizado de 0,001, momentum de 0,9, por 10 épocas com tamanho de lote de 32. A boa performance do modelo CNN em automação de semicondutores e a capacidade de representar características complementares diferem-no de abordagens clássicas. A arquitetura da CNN pode ser vista na Figura 6.

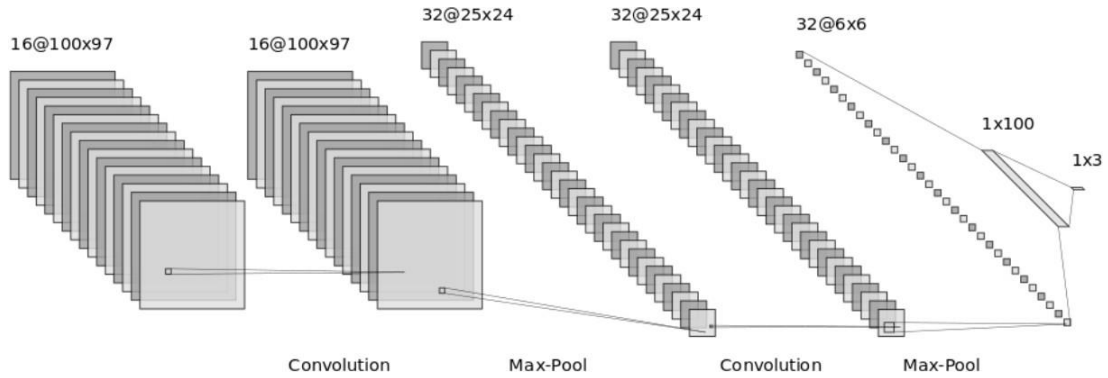


Figura 6. Arquitetura CNN proposta

O treinamento do modelo de classificação é realizado a partir de um conjunto de imagens de entrada e a fase de testes diz respeito a outro conjunto de imagens de entrada. A técnica de validação cruzada *k-fold*, sendo $k=5$ foi utilizada. O conjunto de dados utilizado para o processo de treinamento é composto por 2913 imagens das categorias de esferas de solda, adquiridas após o processo de detecção das esferas.

3.4 Métricas de Avaliação

A avaliação da arquitetura envolve a utilização da métrica da Interseção sobre a União (IoU, do inglês, *Intersection over Union*), que calcula a interseção entre as caixas delimitadoras previstas e as reais fornecidas durante a rotulagem. A Figura 7 fornece uma explicação intuitiva sobre como esse cálculo é realizado. Além disso, um limiar L é definido para determinar se a inferência é classificada como Verdadeiro Positivo (VP) - quando o IoU é maior ou igual a L - Falso Positivo (FP) - quando o IoU é menor que L - ou Falso Negativo (FN) - quando não há previsão para essa área na classe correspondente. A partir dessas detecções, são calculados o *recall* (Equação 1), a precisão (Equação 2) e a *Average Precision* (AP) para cada classe, além da média dessas APs para o limiar L , conhecida como *mean Average Precision* (mAP). Adicionalmente, é calculada a área sob a curva de Precisão x *Recall* para cada classe (Padilla et al, 2020).

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{\text{Área de Sobreposição}}{\text{Área da União}}$$

Figura 7. Cálculo da Interseção sobre a União (IoU) dada a caixa real e predita. Fonte: Padilla et al. (2020)

$$Recall = \frac{VP}{VP + FN} \quad (1)$$

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (2)$$

4. Resultados e Discussões

Nesta seção, apresentamos os resultados e discussões referentes ao modelo treinado a partir do algoritmo YOLOv5, utilizando imagens e anotações da base de referência. A seção está dividida em duas partes: (i) discussão das métricas de avaliação de desempenho do modelo e (ii) comparação da classificação obtida por meio do modelo de CNN com outras abordagens.

4.1 Avaliação das Detecções de Esferas de Solda (YOLOv5)

Os resultados obtidos para o conjunto de treinamento e validação são expostos na Tabela 3. Observando os resultados da tabela, é possível verificar o bom desempenho da arquitetura escolhida para a problemática levantada ao longo deste trabalho. A média dos valores de $mAP@50$ ficaram em 92,6%.

Tabela 3. Métricas obtidas ao final das 300 épocas de treinamento.

Classe	Precisão	Recall	$mAP@50$	$mAP@50:95:100$
Todas	83,6%	90,1%	92,6%	65,7%
Ausente	81%	91,6%	90,6%	63,3%
Correta	87,6%	90,6%	95,7%	67,7%
Falha	82,1%	88,3%	91,5%	66,1%

As métricas relevantes são mostradas na Figura 8. O processo de treinamento mostrou tendência decrescente de pontuação de perda de caixa normalizada ao longo das épocas para o conjunto de treinamento. A função de perda de caixa representa a precisão com que o algoritmo pode localizar o centro de um objeto (categoria de solda) e o quão bem a caixa delimitadora prevista cobre um objeto. A perda de objetividade (obj_loss) é uma medida da probabilidade de que um objeto exista em uma região proposta de interesse.

A Figura 8 também ilustra a tendência da pontuação de perda de caixa normalizada ao longo dos períodos de treinamento. A perda de caixa diminui ao longo da fase de treinamento no conjunto de treinamento e validação. Isso implica em bom potencial de generalização do modelo YOLOv5s na localização de esfera de solda, falha de solda ou ausência de solda. A perda de objetividade no conjunto de validação apresentou tendência ascendente após 100 épocas, portanto o treinamento foi interrompido precocemente.

Os escores de precisão e *recall* diminuíram consistentemente durante o treinamento. Mais imagens no conjunto de dados podem potencialmente levar ao aumento da pontuação de *recall*. É possível notar como a pontuação de precisão e *recall* aumentaram ao longo das épocas de treinamento. O modelo treinado estava detectando as categorias de solda com precisão considerável.

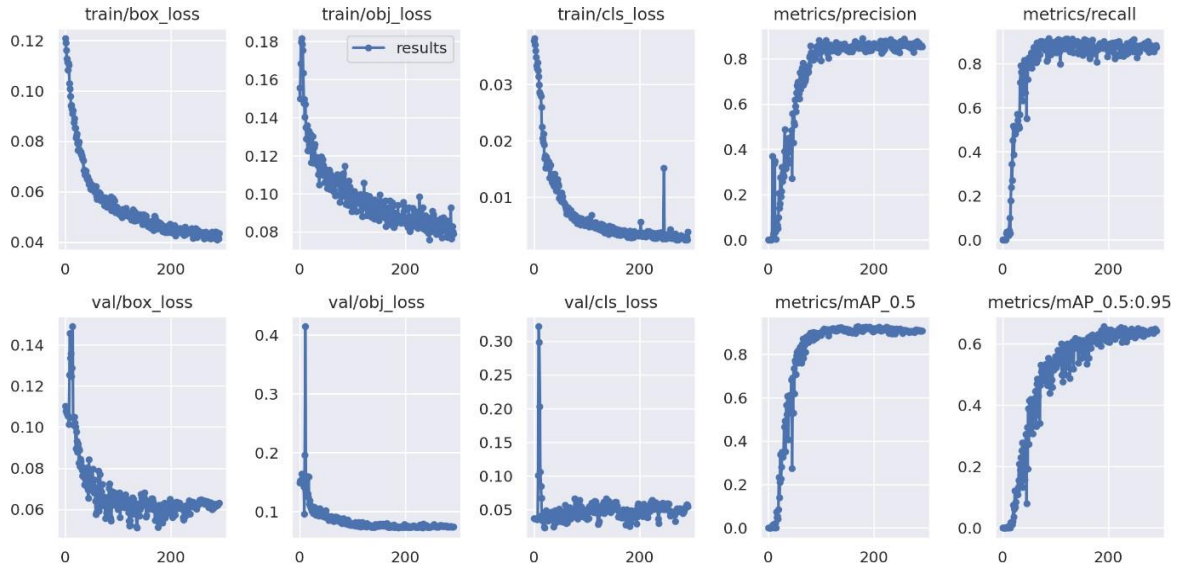


Figura 8. Visão geral dos resultados do treinamento do modelo baseado no detector YOLOv5s. Os gráficos ilustram a curva da função de perda para três diferentes variáveis, considerando tanto o conjunto de treinamento (train), quanto o conjunto de validação (val) e os comportamentos da precisão e recall e os resultados de mAP do modelo

A Figura 9 sintetiza o resultado das classificações para cada categoria por meio de uma matriz de confusão. Através dela observa-se que as categorias foram detectadas com uma taxa considerável de precisão.

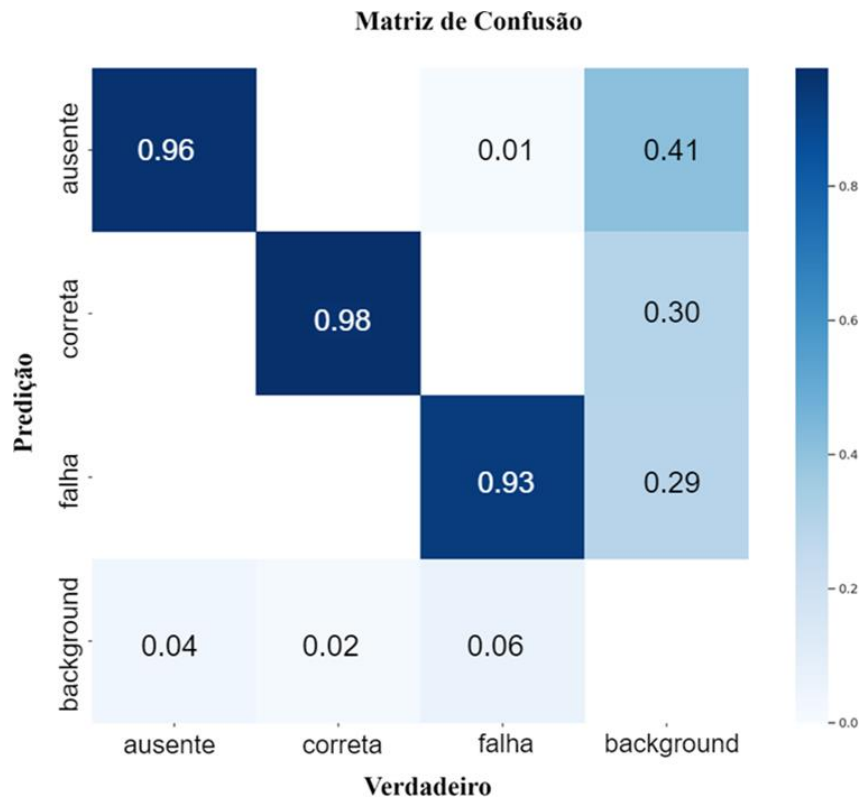


Figura 9. Matriz de confusão das detecções obtidas no conjunto de treinamento

A fim de ilustrar o desempenho do modelo, uma comparação entre as imagens de treinamento e teste é mostrada na Figura 9. Esta figura compara os dados utilizados no treinamento do modelo, manualmente rotulados, com as previsões desempenhadas nestes mesmos dados. Observa-se que o modelo detectou as três categorias propostas neste estudo. Tal desempenho retrata a capacidade do modelo de aprender, com relativa eficácia, os padrões necessários para caracterizar uma esfera de solda com base no conjunto de dados fornecidos, justificando a precisão e o *recall* obtidos durante o treinamento.

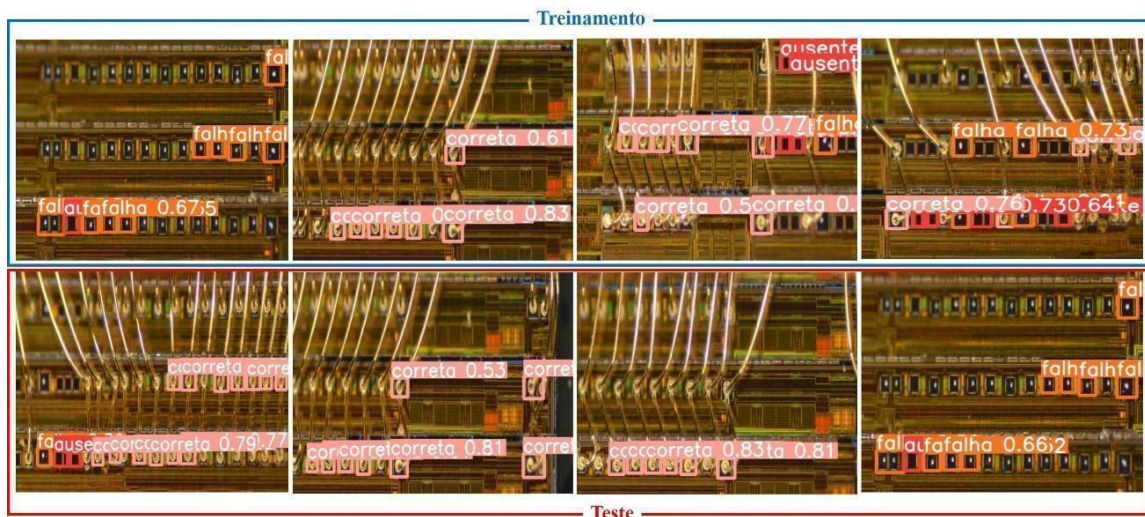


Figura 9. Performance do modelo nos dados utilizados para o treinamento e teste

No entanto, o objetivo do modelo é desempenhar essa mesma performance em dados nunca vistos anteriormente, generalizando e perpetuando o aprendizado. Para avaliar esse comportamento foi utilizado um conjunto de teste para previsões. As imagens utilizadas nesse teste foram obtidas do mesmo modo que as imagens utilizadas no treinamento e validação. São imagens diferentes, que não sofreram rotulação prévia e não foram vistas pelo modelo até então. Por fim, a Figura 10 ilustra imagens pertencentes ao conjunto de teste que foram submetidas ao processo de inferência para detecção da condição da esfera de solda localizados na pastilha de silício.

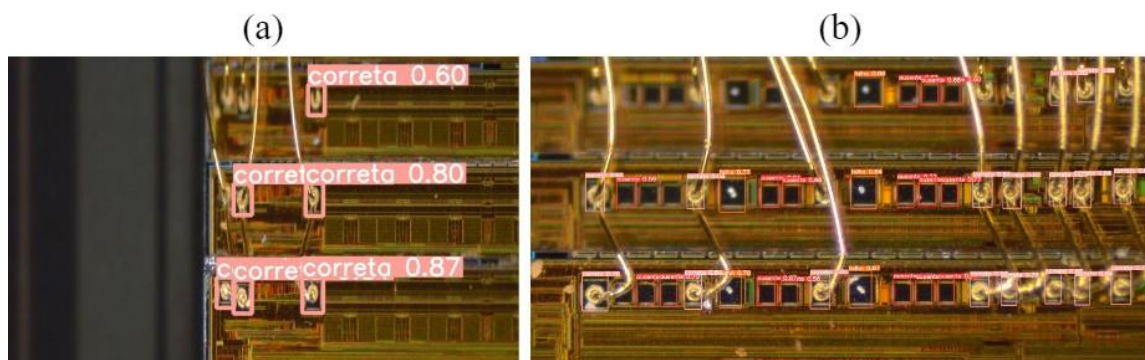


Figura 10. Exemplos de categorias de solda detectados após o treinamento do modelo. Na figura (a) identificação da presença de categoria de solda correta. Na figura (b) detecções das três categorias de solda

4.2 Avaliação das Classificações de Esferas de Solda (CNN)

Tanto o modelo CNN quanto o modelo YOLO foram treinados, validados e testados na mesma base de imagens. No processo de classificação, verifica-se que o modelo CNN obteve uma acurácia média de 97,87% com um desvio padrão de ± 0.560 .

Diferentes abordagens para lidar com o problema de classificação de esferas de solda foram implementadas e avaliadas. As abordagens de comparação são: i) Local Binary Pattern (LBP) e classificador SVM, com kernel polinomial de terceiro grau; ii) classificador LBP e Ada Boosting (AB); iii) classificador LBP e Random Forest (RF); iv) Histograma de Gradiente Orientado (HOG) e classificador SVM, com núcleo polinomial de terceiro grau; v) classificador HOG e AB; e vi) classificador HOG e RF. Essas técnicas de comparação foram utilizadas devido aos bons resultados obtidos na análise automática de semicondutores (Iglesias. et al., 2021) e no contexto geral de inspeção visual automática (Rahman et al., 2019) (Thielen et al., 2020). Os resultados alcançados mostram que o modelo CNN proposto supera as outras técnicas clássicas, como podemos observar na Tabela 4.

Tabela 4. Resultados da classificação da esfera de solda.

Método	Acurácia
LBP + SVM Poly 3	26.208 \pm 3.345
LBP + AB	84.345 \pm 2.937
LBP + RF	90.627 \pm 0.975
HOG + SVM Poly 3	19.342 \pm 3.076
HOG + AB	87.241 \pm 2.088
HOG + RF	95.114 \pm 0.736
Método Proposto	97,87 \pm 0.560

5. Conclusão

O trabalho em questão abordou o problema de inspeção visual de soldagem no encapsulamento de semicondutores. A principal contribuição do presente trabalho foi fornecer uma abordagem para detectar e classificar categorias de esferas de solda de três classes de objetos. Ao contrário de outras metodologias, o método proposto atinge alta precisão e apresenta grande capacidade de inspecionar condições de esferas de solda muito pequenas, proporcionando melhoria no processo de encapsulamento e produção de semicondutores.

O modelo de rede neural YOLOv5, para a tarefa de detecção, alcançou precisão de 83,6%, *recall* de 90,1% e *mAP@50* de 92,6%, e o modelo de CNN, para a tarefa de classificação, obteve uma acurácia média de 97,87%. Esses resultados são considerados

satisfatórios, considerando a base de imagens utilizada nos experimentos. As imagens das pastilhas de silício continham muitos objetos com propriedades de cor ou textura semelhantes às categorias de soldas propostas neste estudo.

A partir dos resultados obtidos, verificou-se que o modelo treinado alcançou notável capacidade de generalização frente às categorias de objetos proposta neste estudo. Como limitações, não foi possível realizar uma comparação justa para o desempenho do modelo YOLOv5 frente à outros modelos de detecção. Por fim, nessa linha de pesquisa, trabalhos futuros podem ser realizados tomando como ponto de partida o estudo comparativo realizado com a arquitetura YOLO e a CNN. Também pretende-se concentrar esforços para estender a abordagem de inspeção visual automática para lidar com outros tipos de problemas relacionados ao encapsulamento de semicondutores. A análise e inspeção das ligações dos fios também é um problema relevante para investigação e se faz necessário incorporar nas linhas de produção. O repositório deste trabalho juntamente o *dataset* utilizado encontra-se disponível nas referências.

Referências

- Alexandrova, S., Tatlock, Z., and Cakmak, M. (2015). **Roboflow: A flow-based visual programming language for mobile manipulation tasks**. In 2015 IEEE International Conference on Robotics and Automation (ICRA), pages 5537–5544. IEEE.
- Brosnan, T. and Sun, D.-W. (2004). **Improving quality inspection of food products by computer vision—a review**. Journal of food engineering, 61(1):3–16.
- Cao, Y. (2021). **Design and application of robot vision inspection system for semiconductor metal target**. In 2021 3rd International Symposium on Robotics Intelligent Manufacturing Technology (ISRIMT), pages 464–467.
- Cao, Z., Simon, T., Wei, S.-E., and Sheikh, Y. (2017). **Realtime multi-person 2d pose estimation using part affinity fields**. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7291–7299.
- Duffy, G. L. (2013). **The ASQ quality improvement pocket guide: basic history, concepts, tools, and relationships**. Quality Press.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016) **Deep learning**. MIT press.
- HAYKIN, Simon. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2001.
- Hartmann, B. (2014). **Human worker activity recognition in industrial environments**. KITScientific Publishing.
- Iglesias., B., Otani., M., and Oliveira., F. (2021). **Glue level estimation through automatic visual inspection in pcb manufacturing**. In Proceedings of the 18th International Conference on Informatics in Control, Automation and Robotics - ICINCO, pages 731–738. INSTICC, SciTePress.
- Jia Yao 1, 2. J. Q. J. Z. H. S. 1. ,. J. Y. a. X. L. 1., s.d. **A Real-Time Detection Algorithm for Kiwifruit Defects Based on YOLOv5**. Em: s.l.:MDPI.
- Jin, R. & Niu, Q., 2021. **Automatic Fabric Defect Detection Based on an Improved YOLOv5**. Em: Mathematical Problems in Engineering. s.l.:s.n.
- Jocher, G., Nishimura, K., Mineeva, T., and Vilarino, R. (2020a). **Yolov5 (2020)**. **GitHub**

- repository:** <https://github.com/ultralytics/yolov5>.
- Khan, A., Sohail, A., Zahoor, U., and Qureshi, A. S. (2020). **A survey of the recent architectures of deep convolutional neural networks**. *Artificial intelligence review*, 53(8):5455–5516.
- Klco, P., Koniar, D., Hargas, L., and Paskala, M. (2022). **Automated detection of soldering splashes using yolov5 algorithm**. In 2022 International Conference on Applied Electronics (AE), pages 1–4
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). **Gradient-based learning applied to document recognition**. *Proceedings of the IEEE*, 86(11):2278–2324.
- Lee, Y. T., Kumaraguru, S., Jain, S., Robinson, S., Helu, M., Hatim, Q. Y., Rachuri, S., Dornfeld, D., Saldana, C. J., and Kumara, S. (2017). **A classification scheme for smart manufacturing systems' performance metrics**. *Smart and sustainable manufacturing systems*, 1(1):52.
- Li, X., Tian, M., Kong, S., Wu, L., and Yu, J. (2020). **A modified yolov3 detection method for vision-based water surface garbage capture robot**. *International Journal of Advanced Robotic Systems*, 17(3):1729881420932715.
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). **Focal loss for dense object detection**. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., and Berg, A. C. (2016). **Ssd: Single shot multibox detector**. In *European conference on computer vision*, pages 21–37. Springer.
- Malta, A., Mendes, M. & Farinha, T., 2021. **Augmented Reality Maintenance Assistant Using YOLOv5**. MDPI -Applied Sciences.
- Padilla, R., Netto, S. L., and da Silva, E. A. B. (2020). **A Survey on Performance Metrics for Object-Detection Algorithms**. In 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), pages 237–242, Niterói, Brasil.
- PEREIRA, P. V. L. **Página do repositório do código deste projeto**. 01/2022. Disponível em: <https://github.com/paulovitor21/tcc-soldering-balls>.
- Rahman, N. N. S. A., Saad, N. M., Abdullah, A. R., and Ahmat, N. (2019). **A Review of Vision Based Defect Detection Using Image Processing Techniques for Beverage Manufacturing Industry**. *Jurnal Teknologi*, 81(3).
- Silva, C., Ferreira, N., Meireles, S., Otani, M., Silva, V., Freitas, C., and Oliveira, F. (2022). **The visual inspection of solder balls in semiconductor encapsulation**. Pages 750–757.
- Srivastava, A., Tolle, I., Mraz, A., Gupta, S., Huang, R., Nguyen, H., Dey, L., Jain, A., Lee, S.-H., and Kini, S. (2016). **Non-traditional inspection strategy for inline monitoring in excursion scenarios: Defect inspection**. In 2016 27th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), pages 197–200
- Thielen, N., Werner, D., Schmidt, K., Seidel, R., Reinhardt, A., and Franke, J. (2020). **A machine learning based approach to detect false calls in smt manufacturing**. In 2020 43rd International Spring Seminar on Electronics Technology (ISSE), pages 1–6.

- YAMASHITA, Rikiya et al. **Convolutional neural networks: an overview and application in radiology**. *Insights Into Imaging*, v. 9, n. 4, p.611-629, 22 jun. 2018.
- Zhang, K., Huang, T., Su, Z., and Guan, T. (2021). **Design of solder quality inspection system based on machine vision**. In 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), volume 5, pages 2413–2418.
- Zhao, Z.-Q., Zheng, P., Xu, S.-t., and Wu, X. (2019). **Object detection with deep learning: A review**. *IEEE transactions on neural networks and learning systems*, 30(11):3212– 3232.
- Zhao, Z., Zheng, P., Xu, S. & Wu, X., 2019. **Object detection with deep learning**. *IEEE Trans. Neural Netw.*
- Zhou, D., Zhang, G., and Guo, Y. (2017). **Detecting multi-interface from oil-water separation image based on difference statistics method**. In 2017 IEEE International Conference on Smart Cloud (SmartCloud), pages 191–196.
- Zhou, J. G., Chen, H., Long, Y., Wang, K., Guo, H., and Liu, F. (2021). **Backside defect monitoring strategy and improvement in the advanced semiconductor manufacturing**. In 2021 China Semiconductor Technology International Conference (CSTIC), pages 1–5.