

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

Paulo Vítor Libório Pereira¹, Felipe Gomes de Oliveira¹

¹Instituto de Ciências Exatas e Tecnologia – Universidade Federal do Amazonas (UFAM) – Itacoatiara – AM – Brasil

vitor.liborio@icomp.ufam.edu.br, felipeoliveira@ufam.edu.br

Resumo. *A detecção de objetos baseada em Visão Computacional é essencial para acelerar a produção de produtos eletrônicos. Para o treinamento e teste da rede neural, foi utilizada a métrica mAP para avaliar os resultados da detecção 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 devido a dimensão dos componentes ser muito reduzida, 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 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 onde 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 (Silva et al., 2022).

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 classificação das esferas de solda contidas nas caixas delimitadoras encontradas.

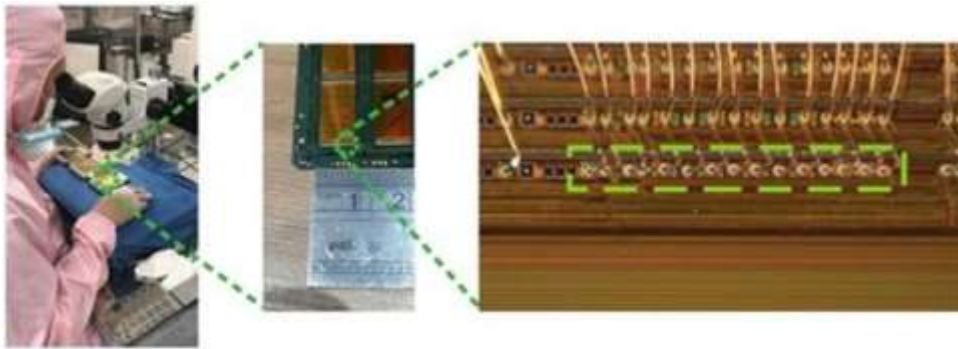


Figura 1. Inspeção visual convencional no processo de soldagem de pastilha de silício.

Este estudo se propõe a colaborar na resolução do problema de inspeção automática para o 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 (You Only Look Once); e *ii*) na classificação de condições de solda, usando um modelo de Rede Neural Convolutiva (CNN).

2. Fundamentação Teórica

2.1 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNNs), descritas por (LeCun et al., 2001), são arquiteturas comuns de Aprendizado Profundo utilizadas na resolução de problemas de Visão Computacional, especialmente para classificação de imagens. Em resumo, CNNs são redes neurais que aplicam, pelo menos em uma camada, uma operação chamada Convolução para aprender representações em imagens (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).

2.2 Detecção de Objetos com YOLOv5

Nos últimos anos, a detecção de objetos com base em aprendizado profundo tem sido uma linha 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 CNNs.

Dentre os diversos algoritmos existentes para detecção e reconhecimento de imagens, os algoritmos R-CNN e YOLO são os mais destacados. Embora o algoritmo R-CNN tenha melhor acurácia, sua velocidade de detecção é muito mais baixa do que a apresentada pelo algoritmo YOLOv5, o que o torna inapropriado 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. O 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 do 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 possível saída.

3. Metodologia

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 de Imagens; *iii*) Detecção das Esferas de Solda; e *iv*) Classificação das Esferas de Solda. 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.

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 subconjunto de treinamento, um subconjunto de validação e um subconjunto de teste (não havendo intersecção entre os subconjuntos). Os subconjuntos de treinamento e validação foram utilizados para o treino e o subconjunto de teste foi usado para avaliação do modelo. Para o processo de avaliação mencionado foi utilizado o protocolo validação cruzada, considerando o processo *k-fold*, com o valor de *k* igual a 5, ou seja, com a proporção 80% de dados para treinamento e 20% de dados para teste.

3.1 Aquisição de Imagens

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 do Pólo Industrial de Manaus.

3.2 Pré-Processamento de Imagens

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 disponíveis no repositório (Jocher et al., 2020). Em seguida, foi utilizada a técnica de aumento de dados (*data augmentation*), usando a ferramenta Roboflow (Alexandrova et al. 2015). 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. Em seguida, as imagens foram exportadas no formato YOLOv5 PyTorch.

3.3 Detecção das Esferas de Solda

Nesta etapa é realizado o processo de detecção das esferas de solda presentes nas imagens a serem inspecionadas. Para isso, foi utilizado o algoritmo YOLO em sua quinta versão. O YOLOv5 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.

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 na imagem e na localização do objeto de interesse.

3.4 Classificação das Esferas de Solda

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). Esse modelo usa ReLU como função de ativação, Max Pooling de tamanho (2,2), e foi treinado com o algoritmo de otimização 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.

4. Resultados e Discussões

Nesta seção, apresentamos os resultados e discussões referentes ao problema de inspeção de esferas de solda no encapsulamento de semicondutores. A seção está dividida em duas partes: *i*) avaliação das técnicas de detecção; e *ii*) comparação da classificação obtida por meio do modelo de CNN com outras abordagens.

4.1. Avaliação das Detecções

Os resultados obtidos para o conjunto de treinamento e validação são expostos na Tabela 1. Observando os resultados da tabela, é possível verificar o bom desempenho da arquitetura escolhida para o problema de detecção de esferas de solda levantado ao longo deste trabalho. A média dos valores de $mAP@50$ foi de 92,6% de acurácia nas detecções de esferas de solda.

Tabela 1. 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%

A Figura 2 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.

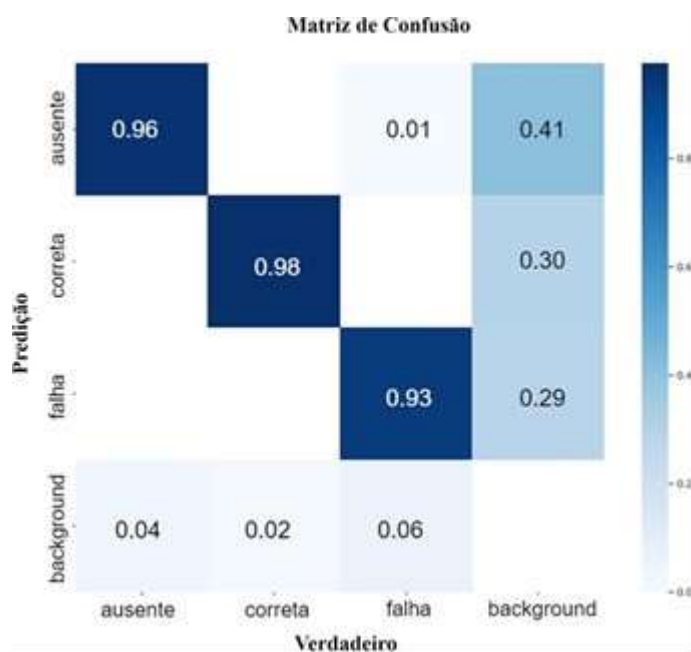


Figura 2. Matriz de confusão das detecções obtidas no conjunto de treinamento.

4.2 Avaliação das Classificações de Esfera 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*) LBP e classificador Ada Boosting (AB); *iii*) LBP e classificador Random Forest (RF); *iv*) Histograma de Gradiente Orientado (HOG) e classificador SVM, com kernel polinomial de terceiro grau; *v*) HOG e classificador AB; e *vi*) HOG e classificador 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 2. Mais detalhes podem ser encontrados no artigo (Silva et al., 2022) e no capítulo de livro (Pereira et al. 2023).

Tabela 2. 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

6. 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. 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 a 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.

Referências

- Alexandrova, S., Tatlock, Z., and Cakmak, M. (2015). Roboflow: A flow-based visual programming language for mobile manipulation tasks. In 2015 IEEE Int. Conference on Robotics and Automation (ICRA), pages 5537–5544. IEEE.
- 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>.
- 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.
- Malta, A., Mendes, M. & Farinha, T., (2021). Augmented Reality Maintenance Assistant Using YOLOv5. MDPI -Applied Sciences.
- Pereira, P. V. L., Silva, C. N., Ferreira, N. P., Meireles, S. S., Otani, M., da Silva, V. J., de Freitas, C. A. O., and Oliveira, F. G. (2023). Automatic fault detection in soldering process during semiconductor encapsulation. In Informatics in Control, Automation and Robotics, pages 130–145, Cham. Springer International Publishing.
- 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. In Proceedings of the 19th International Conference on Informatics in Control, Automation and Robotics - ICINCO, pages 750–757. INSTICC, SciTePress.
- Yamashita, Rikiya et al. (2018) Convolutional neural networks: an overview and application in radiology. Insights Into Imaging, v. 9, n. 4, p.611-629.
- Zhang, K., Huang, T., Su, Z., and Guan, T. (2021). Design of solder quality inspection system based on machine vision. In 2021 IEEE 5th Adv. Info. Tech., Elect. 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 trans. on neural networks and learning systems.