Cody Costa

9/20/2025

EE263

Wafer Image Classification Using Deep Learning CNN Architecture

Abstract:

Wafer image classification plays a critical role in the semiconductor manufacturing industry, where the early detection of defects directly impacts yield, reliability, and production costs. In this work, a convolutional neural network (CNN) was developed to automatically classify wafer images into distinct categories based on their visual features. A dataset of 26×26-pixel wafer images was pre-processed, normalized, and augmented with label encoding to prepare it for supervised training. The CNN architecture combined convolutional, pooling, and fully connected layers with dropout regularization, and was trained using the Adam optimizer with categorical cross-entropy loss. Experimental evaluation demonstrated the model’s ability to achieve high classification accuracy on unseen test data, highlighting the effectiveness of deep learning for defect detection tasks.

Methodology:

The following high level abstracted methodology was adopted to develop and evaluate the CNN for wafer image classification.

**1. Dataset Preparation**

A wafer image dataset stored in a serialized .pkl file was used as the input source. The dataset contained both image data and their corresponding labels. Depending on the format (Pandas DataFrame or dictionary), the image arrays and labels were extracted accordingly.

All images were pre-processed to ensure consistency in dimensionality and format. Specifically:

* Images originally in channel-first format (C × H × W) were converted to channel-last format (H × W × C).
* Grayscale images were converted to three-channel RGB images using OpenCV.
* All images were resized to a uniform dimension of 26 × 26 pixels.
* Pixel intensities were normalized to the range [0,1] to improve training stability.

**2. Train/Test Split**

The dataset was divided into training and testing subsets using an 80:20 ratio. A fixed random seed (42) was applied to ensure reproducibility of the split. Within the training set, a validation split of 10% was further used during model training to monitor generalization performance.

**3. Model Architecture**

Our CNN was designed using the Keras Sequential API. The architecture consisted of:

* **Convolutional layers:** Two convolutional layers with 32 and 64 filters respectively, both using **3×3 kernels** and ReLU activation.
* **Pooling layers:** Each convolutional layer was followed by 2×2 max pooling for spatial dimensionality reduction.
* **Fully connected layers:** The flattened output was passed to a dense layer with 128 neurons and ReLU activation.
* **Dropout regularization:** A dropout rate of 0.5 was applied to reduce overfitting.
* **Output layer:** A dense softmax layer with units equal to the number of classes provided with class probability predictions.

A graphical model representation is as follows:

A screenshot of a computer

AI-generated content may be incorrect.

**4. Model Training**

The CNN was trained for 20 epochs with a batch size of 32.

Experimental Results:

Our CNN trained on wafer images achieved strong and stable performance across 20 epochs. Here is a quick summary of the accuracy benchmarks:

* **Training Accuracy:** Gradually improved from 93.7% in epoch 1 to 94.6% by epoch 20.
* **Validation Accuracy:** Remained consistently high, stabilizing around 93.8–94.2%, indicating good generalization without significant overfitting.
* **Training Loss:** Decreased steadily from 0.4217 to 0.2333, reflecting effective optimization.
* **Validation Loss:** Declined from 0.3455 to 0.2475, supporting stable validation performance.
* **Final Test Accuracy**: 94.71%, with a test loss of 0.2272, confirming that the model generalized well to unseen data.

A screen shot of a computer

AI-generated content may be incorrect.

Conclusions:

The CNN demonstrates excellent classification performance on wafer images, achieving nearly 95% accuracy while maintaining low loss. These results suggest the model is highly effective for automated defect detection and classification in semiconductor manufacturing applications.