

ВИКОРИСТАННЯ *MASK R-CNN* ДЛЯ АВТОМАТИЗОВАНОГО ДЕТЕКТУВАННЯ ДЕФЕКТІВ ВИРОБІВ ЗІ СКЛА

Привала С. В., Сазонов А. Ю.

КПІ ім. Ігоря Сікорського, s.pryvala@gmail.com

Процес автоматизованого детектування дефектів у виробках зі скла за змістом є послідовністю етапів, на кожному з яких виконується низка операцій, що націлені на отримання якісного результату. Взаємозв'язок етапів можна представити у вигляді послідовного з'єднання блоків, що відображено на відповідній структурній схемі на рис. 1.



Рис. 1. Структурна схема роботи системи детектування дефектів

На *етапі 1* відбувається реєстрування виробу системою машинного зору.

Тут під час переміщення об'єкта на рухомому конвеєрі здійснюється його фотофіксація. Кожне зображення зберігається у базі даних для подальшого аналізу разом із службовою інформацією про об'єкт.

На *етапі 2* виконується попередня підготовка зображень до наступної обробки, яка включає:

- фільтрацію зображення для зменшення шуму та підвищення якості роботи системи;
- масштабування зображення до розмірів, що вимагаються системою візуальної обробки.

На *етапі 3* виконується безпосередня обробка зображення модулем автоматизованого детектування дефектів. У результаті отримуємо усі наявні класи, їх положення. У подальшому, для більш точної класифікації, обчислюються лінійні розміри кожного дефекту та розбиваються на класи та підкласи відповідно до розміру.

Під поняттям клас розуміється група, до якої можна віднести конкретний дефект, наприклад, “подряпина”, “мошка”, “тріщина”, тощо.

Для обробки та аналізу отриманих графічних даних запропоновано використання архітектури штучної нейронної мережі (ШНМ) під назвою *Mask R-CNN*, що має ряд переваг у порівнянні із іншими:

1. Така архітектура ШНМ є найточнішою у класі систем із сегментації дефектів і є яскравим прикладом загорткової нейромережі (*Convolutional Neural Network – CNN*);

2. Запропонована ШНМ може бути навчена самостійно на власних даних, оскільки вона побудована за модульним принципом і не потребує великих апаратних потужностей. Найскладнішим етапом багатьох реалізацій систем машинного зору на основі ШНМ є підготовка даних для їх навчання, що є непростою задачею та часто вимагає спеціальних програмних рішень. Для запропонованої реалізації підготовка даних проводилась шляхом розмітки, тобто вказувалось, де на даних навчальної множини зображень знаходиться той чи інший клас (дефект), та готувався його програмний опис, відповідно до прикладу;

3. Простота роботи. До переваг можна віднести простоту її структури та модульність, тобто для специфічних задач, можливо замінити модулі, що змінить складність модель та призведе до можливого покращення результату. Це надасть змогу розробляти системи на основі даної технології для різних цілей. Наприклад, для розпізнавання у реальному часі необхідними є простіші компоненти (моделі), які розраховуються швидше, але, можливо, не так точно в залежності від задачі. В іншому випадку, у разі заміни легких моделей на більш складні зможемо досягти кращих результатів за її тривалішої роботи.

Архітектура *Mask R-CNN* є розвитком попередньої *R-CNN* і працює таким чином – оригінальний алгоритм *R-CNN* складається з 4-х кроків:

1. Вхідне зображення подається у мережу;
2. Розрахунок потенційних регіонів (тобто області зображення, які містять чи можуть містити об'єкти) за допомогою алгоритму селективного пошуку;
3. Використання *Transfer Learning (TL)*, а саме виокремлення особливостей зображення для подальшого обчислення особливостей кожної пропозиції (які фактично є областями інтересів – *ROI – Regions of Interests*) із використанням попередньо навченої *CNN*. *TL* дозволяє використовувати накопичений під час розв'язання однієї задачі досвід для вирішення іншої – аналогічної – проблеми;
4. За допомогою методу опорних векторів (*Support Vector Machine – SVM*) на основі оцінки кожній області інтересу присвоюється певний клас.

Цей метод дозволяє отримати надійні розрізнявальні особливості, яким навчена *CNN*.

У *Mask R-CNN* запропоновано два основні вдосконалення:

- заміна модуля об'єднання областей інтересу (*ROI Pooling*) модулем узгодження областей інтересу (*ROI Align*);
- додавання окремої гілки із модулем узгодження областей інтересу (див. рис. 2).

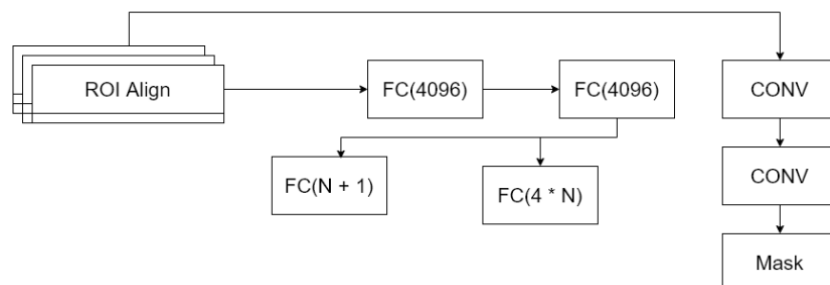


Рис. 2. Структурна схема ШНМ *Mask R-CNN*

Mask R-CNN використовує мережу пропозицій регіонів (*Region Proposal Network – RPN*) для створення областей зображення, які потенційно містять об'єкт.

Кожен із цих «регіонів» класифікується на основі «оцінки об'єктності» (тобто наскільки ймовірно, що дана область потенційно може містити об'єкт), а потім зберігаються вершини N найбільш імовірних областей.

Авторами [1] пропонується використовувати $N = 300$ (кількість областей).

Кожна із 300 обраних областей проходить через три паралельні «гілки» мережі (рис. 2):

- прогнозування міток;
- прогнозування меж об'єкта;
- маска об'єкта.

Під час прогнозування кожна із N областей проходить заглушення немаксимумів (*non-maxima suppression*), та виконується відбір 100 найкращих областей, що призводить до формування $4D$ -тензора $100 \times L \times a \times b$, де L – кількість міток класу в наборі даних, а $a \times b$ – розмір (наприклад, 15×15 пікс.) кожної із масок L .

Сумарна маска містить лише 15×15 пікселів, тому необхідним є відновлення розмірів маски відповідно до розмірності вхідного зображення, після чого змінена розмірна маска може бути накладена на вихідне зображення

Результатом роботи ШНМ є клас дефекту, координати його розташування та маска [2].

1. Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick. Mask R-CNN [Electronic resource]. *Proceedings. Mode of access*. URL: <https://arxiv.org/abs/1703.06870> (Last accessed: 21.03.2020).

2. Rosebrock A. Mask R-CNN with OpenCV [Electronic resource]. *Proceedings. Mode of access*. URL: <https://www.pyimagesearch.com/2018/11/19/mask-r-cnn-with-opencv/> (Last accessed: 21.03.2020).