

A Research on Cylindrical Pill Bottle Recognition with YOLOv8 and ORB

Dae-Hyun Kim*, Hyo Hyun Choi*

*Student, Dept. of Computer Science and Engineering, Inha Technical College, Incheon, Korea

*Professor, Dept. of Computer Science and Engineering, Inha Technical College, Incheon, Korea

[Abstract]

This paper introduces a method for generating model images that can identify specific cylindrical medicine containers in videos and investigates data collection techniques. Previous research had separated object detection from specific object recognition, making it challenging to apply automated image stitching. A significant issue was that the coordinate-based object detection method included extraneous information from outside the object area during the image stitching process. To overcome these challenges, this study applies the newly released YOLOv8 (You Only Look Once) segmentation technique to vertically rotating pill bottles video and employs the ORB (Oriented FAST and Rotated BRIEF) feature matching algorithm to automate model image generation. The research findings demonstrate that applying segmentation techniques improves recognition accuracy when identifying specific pill bottles. The model images created with the feature matching algorithm could accurately identify the specific pill bottles.

▶ **Key words:** Image Detection, Feature Matching, Image Stitching, YOLOv8, Image Segmentation

[요 약]

본 논문은 영상에서 특정 원통형 약통을 식별할 수 있는 모델 이미지 생성 방식을 제시하고 데이터 수집에 대한 기술을 연구한다. 기존 연구들은 객체 인식과 특정 객체 식별이 분리되어 있어 이미지 스티칭(image stitching) 자동화에 적용하기 어려웠으며, 좌표 기반 이미지 추출 방식이 이미지 스티칭 과정에서 객체 영역 외의 정보도 모델 이미지에 포함시키는 문제를 갖고 있었다. 이를 해결하기 위해 본 논문은 최근에 출시된 YOLOv8(You Only Look Once)의 세그멘테이션(segmentation) 기법을 수직축 회전하는 약통 영상에 적용하고 특징점 매칭 알고리즘인 ORB (Oriented FAST and Rotated BRIEF)를 활용하여 모델 이미지 생성을 자동화하였다. 연구 결과, 세그멘테이션 기법을 적용할 경우 특정 약통 식별시 인식률이 향상되었으며 특징점 매칭 알고리즘으로 생성된 모델 이미지는 특정 약통을 정확하게 식별해 낼 수 있었다.

▶ **주제어:** 객체 인식, 특징 매칭, 이미지 스티칭, YOLOv8, 이미지 세그멘테이션

-
- First Author: Dae-Hyun Kim, Corresponding Author: Hyo Hyun Choi
 - *Dae-Hyun Kim (rnasterofmysea@gmail.com), Dept. of Computer Science and Engineering, Inha Technical College
 - *Hyo Hyun Choi (hchoi@inhac.ac.kr), Dept. of Computer Science and Engineering, Inha Technical College
 - Received: 2023. 11. 06, Revised: 2024. 01. 02, Accepted: 2024. 01. 08.

I. Introduction

약물 안전 및 관리는 현대 의료 및 제약 산업에서 중요한 측면 중 하나로, 특히 환자 및 의료 전문가들은 정확한 약물 관리를 위해 약물 섭취를 체계적으로 추적한다. 이러한 과정은 환자의 안전과 약물 효과성을 보장하기 위해 필수적이다. 그 때문에 약물 섭취 관리에 컴퓨터 기술을 접목하기 위한 다양한 시도가 활발히 이뤄지고 있다. 이 논문은 개별 약통 인식에 초점을 맞추어 약물 섭취 여부를 관리할 수 있는 시스템에 관한 연구를 진행하였다.

특정 단일 객체에 대한 인식은 광학 문자인식(OCR, Optical Character Recognition), 객체 인식(image detection), 특징점 매칭 등 이미지 데이터를 활용하여 식별해내는 방향으로 많은 연구가 진행되고 있다.

특정 객체 검출에 관한 국내 연구 사례에서는 two-stage detector의 Fast-RCNN(Regions with Convolutional Neural Networks), RCNN 등의 모델과 one-stage detector의 YOLO 기반 모델과의 성능 비교에 많은 연구가 진행되었다 [7, 8]. YOLO 버전이 향상되면서 인식률과 bounding box 적중률뿐만 아니라 세그멘테이션(segmentation), 3D 인식 등 전반적인 성능이 크게 향상하고 있음을 알 수 있고 two-stage 기반의 모델보다 인식 속도가 빠르기 때문에 실시간 영상에서 객체를 검출해 내기 위해 YOLO 모델을 본 연구 모델로 채택하였다.[13] 그러나 객체 인식 모델일 경우 타겟 객체가 약통인지 아닌지 검출하는 데 용이하나 해당 약통의 제품이 어떤 제품인지 판별하는 데 있어 제약사항이 많다. 이를 해결하고자 특징점 매칭 알고리즘을 객체 인식 모델 프로세스에 추가해 약통이라는 객체를 인식한 후 해당 약통의 특징점 분석을 통한 특정 약통 인식이라는 프로세스를 설계하였다.

특징점 매칭을 통해 객체를 인식하는 연구에서는 객체에 대한 전방향적인 정보를 획득하고자 수직축 회전에서의 이미지들을 Shift 알고리즘을 통해 하여 특정 약통 식별을 한 연구 [8]에서는 0도에서 360도 사이에서 각도별 이미지들을 추출하여 n도 간격들의 이미지들을 스티칭했을 때 인식률의 차이를 비교하였다.

해당 연구에서는 위 두 프로세스를 통합하여 특정 객체를 검출할 수 있는 모델 이미지 생성을 구현하고 이를 평가하는 연구를 진행하였다. 본 연구는 사용자가 손으로 약통을 수직 축으로 회전시키는 영상에서 약통 세그먼트를 추출하고 이를 스티칭하여 특정 객체 인식을 위한 이미지 생성하는 과정까지 자동화하였다.

본 논문은 YOLOv8을 사용하여 약통 세그멘테이션을 수행하고, 그 다음 특징점 매칭의 ORB 알고리즘을 활용하

여 특정 약통을 식별 과정을 수행하는 방법을 제안한다. 특징점 매칭 정확도를 향상하기 위해 객체 인식 시 바운딩 박스(bounding box) 추출이 아닌 세그멘테이션을 통해 매칭을 진행하고 ORB 알고리즘을 통해 스티칭할 이미지를 자동으로 군집화한다. 이러한 접근 방식은 카메라를 통해 약통을 빠르게 감지하고, 객체 인식의 인식률을 높일 수 있는 이미지 데이터 수집을 효율적으로 하며, 개별 약통을 식별하여 약물 관리 작업을 자동화하고 개선하는 데 도움이 될 것으로 기대된다.

본 논문은 2장에서 시스템에 활용한 컴퓨터 비전의 핵심 기술과 모델을 소개하고 특정 객체 식별에 관해 사전 연구들을 분석한다. 3장은 제안한 시스템과 테스트 방법에 관해 설명하고 4장에서 각 프로세스에서 연구한 결과를 도출한다. 5장은 제약사항, 개선사항과 향후 연구 방안을 포함하여 결론을 맺는다.

II. Preliminaries

1. Related works

1.1 Object Detection

객체 검출(object detection)이란 이미지 내의 객체 위치를 특정(localization)하고 그 속성(classification)을 부여하는 작업을 말한다. 딥러닝을 활용한 객체 검출 기술은 지역화의 문제와 분류의 문제를 동시에 진행하는 1-stage 검출기와 두 문제를 순차적으로 진행하는 2-Stage 검출기로 구분할 수 있다 [9].

1.2. YOLOv8

YOLO(You Only Look Once)는 워싱턴 대학의 Joseph Redmon과 Ali Farhadi가 개발했으며 2015년에 출시된 YOLO는 속도와 정확성이 균형을 잡힌 모델로 인기를 얻게 되었다. 이후 출시된 버전들은 개선된 속도와 정확성을 보여주고 있으며. 특이한 점은 v6가 v7 이후에 출시되었으며 v6, v7, v8은 속도 개선을 위한 알고리즘과 신경망 구조에서 차이가 있으며 각 모델의 업그레이드가 개별적으로 수행되고 있다.

본 연구는 Ultralytics 사가 2023년 1월 출시한 YOLOv8 버전을 사용했으며, 이 버전은 통합 프레임워크를 지원하여 객체의 분류, 세그먼트화, 인식 등 활용도가 높다. YOLOv8은 YOLOv5의 아키텍처를 보완을 유지한 백본을 사용한다. YOLOv5의 CSPLayer를 C2F 모듈(2개의 convolution이 부분 교차 병목)로 교체하여 인식 정확

도를 향상시켰다. 또한, YOLOv8은 분리된 헤드가 있는 anchor-free 프리 모델을 사용하여 오브젝트니스(objectness), 분류(classification), 회귀(regression) 과정을 독립적으로 처리한다. 이 설계를 통해 각 작업을 분리함으로써 모델의 전반적인 정확도를 개선하였다.

2.1 Feature Matching Algorithm

2.1.1 FAST(Features from Accelerated Segment

2006년 영국 케임브리지 대학의 Edward Rosten과 Tom Drummond [12]가 FAST(Features from Accelerated Segment Test) 알고리즘을 제안하였다. FAST에서는 코너 후보 픽셀(P) 주변에 16개 픽셀로 이루어진 원을 그려 그 원 위에 연속되는 12개(n=12)의 픽셀들의 밝기가 P보다 특정 임계값 이상 밝거나 어두우면 코너로 정의하고 코너의 상태를 3가지로 나뉜다 (1).

$$S_{p \rightarrow x} = \begin{cases} d, & I_{p \rightarrow x} \leq I_p - t & \text{(darker)} \\ s, & I_p - t < I_{p \rightarrow x} < I_p + t & \text{(similar)} \\ b, & I_p + t \leq I_{p \rightarrow x} & \text{(brighter)} \end{cases} \quad (1)$$

이를 이용하여 원주 상의 픽셀들의 밝기분포를 16차원의 ternary 벡터로 표현한다. 그리고 이를 decision tree에 입력하여 코너 점 여부를 분류한다 (2).

$$H(P) = (c + \bar{c}) \log_2(c + \bar{c}) - c \log_2 c - \bar{c} \log_2 \bar{c}$$

where $c = |\{p|K_p \text{ is true}\}|$ (number of corners)
and $\bar{c} = |\{p|K_p \text{ is false}\}|$ (number of non corners) (2)

$$H(P) - H(P_a) - H(P_s) - H(P_b)$$

2.1.2 ORB(Oriented FAST and Rotated BRIEF)

ORB는 특징점 검출 알고리즘으로 FAST, descriptor로 BRIEF(Binary Robust Independent Elementary Features)를 적용한 알고리즘이다 [3]. FAST의 한계인 점을 극복하기 위해 픽셀 밝기와 픽셀 간 거리를 이용해 모멘텀을 구한 후 이를 통해 centroid 값을 구한다. Centroid를 적용하여 스케일(scale)과 회전(rotation)에 대해 특징점에 파라미터를 추가했다 (3).

$$m_{pq} = \sum_{x,y \in B} x^p y^q I(x,y), \quad p, q = \{0, 1\}$$

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (3)$$

$$\theta = \arctan(m_{01}/m_{10})$$

ORB 알고리즘은 Descriptor로 선정한 BRIEF의 회전 성능 저하의 문제를 해결하기 위해 BRIEF를 회전시켜 매칭되는 패턴을 찾는 방법인 r-BRIEF 알고리즘을 제안했다.

2.2 Object Recognition using Feature Matching

카메라를 통해 3차원 물체에 대한 인식 및 식별을 하기 위해 다양한 딥러닝과 머신러닝 알고리즘이 활용되고 있다. 객체의 분류가 아닌 특정 객체에 대한 식별이 필요한 경우 객체 인식 딥러닝 모델과 특징점 추출 기반 알고리즘을 결합해 시스템을 고안하는 방법이 주로 사용되고 있다. 심규호 등 [7]은 2-stage 모델인 mask R-CNN과 Yoloact(You only look at coefficients)을 활용하여 객체 인식 및 segment 추출을 진행하였고 RANSAC (RANdom SAMple Consensus) 알고리즘을 통해 객체의 자세를 추정하였다.

객체 인식을 위해서 특징점 매칭 기술을 접목한 사례 [8]에서는 SIFT 알고리즘을 통해 특징점을 분석을 진행하였으며 원통형 객체를 특정 각도로 수직축 회전하여 이미지를 촬영하여 임의의 객체 영상에서 특징점 매칭의 인식을 실험하였다.

물체 인식을 효율적으로 하기 위하여 물체의 회전에 대하여 일정한 각도마다 생성한 모델들을 이용한 모델 공간을 생성할 수 있다. 수직축을 중심으로 회전하며 생성된 모델들은 회전축 근방에서 왜곡 현상이 적다. 이에 신뢰도가 높은 회전축을 중심으로 각각의 모델에 관심영역 (ROI, Region of Interest)을 지정하고 일정 크기로 정규화한 후 영역을 병합하여 모델 이미지를 생성하였다.

최규식 등 [9]은 서로 다른 시기의 동일한 위치에서 촬영된 파노라마 이미지 쌍의 간판 영역을 찾아내고 동일한 간판을 매칭하는 방법을 제안하였다. 객체 검출 기법과 이미지 지역 특징점 매칭 기법을 활용하여 매칭 과정을 수행하였으며, 실험 결과 YOLOv7과 LoFTR(Local Feature matching with Transformers)를 사용할 때 높은 성능을 보였다. 이 논문에서는 YOLOv7과 Faster-RCNN을 이용하여 간판 객체를 검출하였으며 mAP을 비교했다. 비교 결과 YOLOv7 87.7%, Faster-RCNN이 85.4%으로 YOLOv7를 사용했을 때 속도와 인식을 측면에서 성능이 좋다는 것을 알 수 있다. 특징점 매칭으로 LoFTR을 사용했으며 이 알고리즘은 트랜스포머(Transformer)구조를 이용하여 특징점 추출과 매칭을 동시에 수행한다. 이미지 특징을 추출한 후 트랜스포머의 셀프 어텐션(self-attention)원리를 이용해 특징 간 상관 관계를 학습하고 매칭한다. 특정 약통에 대한 식별을 3가지 단계로 나누어 설계한 논문[14]는 바코드 인식, 텍스트 인식(OCR), 특징점 매칭 순으로 결정 트리를 만들어 사각형 약통에 대

한 식별을 테스트했다. 특징점 매칭으로 사각형 약통을 식별할 경우 가장 매칭의 결과가 좋았으나 실험은 약통의 정면으로만 이루어졌다는 점에서 제약사항이 있다. 또한 바코드 인식에 초점이 맞춰져 있기 때문에 바코드가 보이지 않는 상황에서 인식에 제한 사항이 있다. [11, 14]

위에 언급한 연구들은 특징점 매칭과 객체 인식 시스템이 분리되어 있어 통합적인 프로세스에서의 인식을 테스트의 한계가 있었고, 본 연구에서는 영상에서 자동으로 데이터셋을 형성하고 특징점 매칭을 통해 하나의 단일 데이터 이미지를 생성함으로써 프로세스 통합에 대해 연구하였다. 또한 최신 YOLO의 버전인 v8의 세그먼트 추출 기법을 활용하여 특징점 매칭 시 매칭률을 개선하는 설계를 했다. 이를 통해 본 논문은 단일 이미지 생성의 제약조건이었던 데이터 이미지 추출에 대한 대안을 제시하고 통합 프로세스에서 인식을 테스트하여 자동화 시스템 구축에 방법을 제시하였다.

III. The Proposed Scheme

본 연구의 실험은 Google의 Colab laboratory 환경에서 python으로 진행하였다. 크게 3가지 프로세스를 진행하였으며 객체 인식 모델로 YOLOv8을 이용하여 약통 객체의 영역 및 세그먼트를 추출하였다. Image dataset 을 구성하기 위해 특징점 매칭의 ORB 알고리즘을 활용했으며 수집된 약통 이미지 데이터들을 stitching 하여 하나의 단일 이미지 데이터를 생성했다.

테스트는 두 가지 목적을 가지고 있다. 첫 번째는 프레임 단위로 약통 세그먼트를 추출할 때 추출된 프레임 개수가 인식률에 영향을 주는지 테스트한다. 두 번째는 특징점이 중복되지 않는 부분만 추출하여 stitching한 이미지 데이터와 특정 프레임 단위로 추출한 이미지들을 stitching한 이미지 데이터 간의 인식률을 비교한다. 해당 비교 테스트를 통해 출력 accuracy를 분석하여 특정 객체 식별 여부를 판단할 수 있는 임계값 설정을 최종 목표로 한다.

수직축으로 회전하는 약통을 촬영한 영상에서 YOLOv8에서 Microsoft사의 CoCo(Common Objects in Context) 모델을 활용하여 사전 학습된 모델을 전이 학습하여 세그먼트를 추출한다.

추출된 세그먼트들 집합에서 회전 여부를 파악하기 위해 정면 세그먼트 1개와 n개의 전체 세그먼트 집합과의 특징점 매칭을 진행한다. 각 프레임 간의 객체 회전 연속성과 노이즈 검출을 위해 Affine 함수를 사용한다.

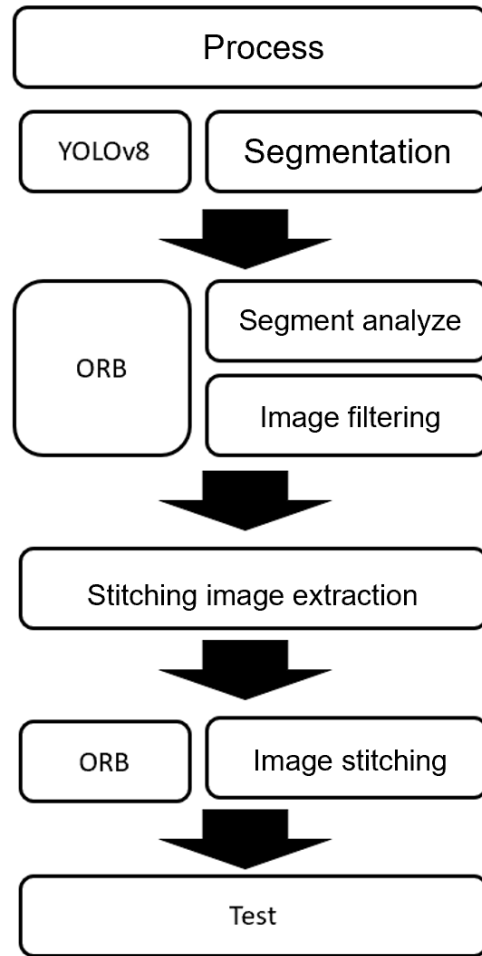


Fig. 1. System process structure



Fig. 2. Yolov8 Pill bottle segmentation

최소자승법을 사용하여 Affine 변환 매개변수를 추정하며 수식 4는 Affine 함수의 계산과정이다. 2×2 행렬 A는 스케일링, 회전, 그리고 기울임에 대한 정보를 담고 있다. 행렬 B는 이동에 대한 이동 벡터이다. 회전성 검출을 위해 현재 프레임과 다음 프레임과 Affine 함수를 활용하여 변환 매개변수(M)을 구한 후 회전과 픽셀 이동을 분석한다.

흔들린 프레임은 라플라시안(Laplacian) 연산자를 통해 제거하였다. 프레임을 라플라시안 연산자의 결과값이 임계값 미만으로 떨어지면 사진이 흔들려서 image stitching을 진행할 수 없는 것으로 간주하고 탈락시킨다.

$$A = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix}_{2 \times 2} \quad B = \begin{bmatrix} b_{00} \\ b_{10} \end{bmatrix}_{2 \times 1} \quad (4)$$

$$M = [A \quad B] = \begin{bmatrix} a_{00} & a_{01} & b_{00} \\ a_{10} & a_{11} & b_{10} \end{bmatrix}_{2 \times 3}$$

필터링 과정을 거친 세그먼트 데이터셋은 수평 보정, 약통의 크기 조정 후 이를 활용하여 이미지 스티칭에 필요한 세그먼트들을 추출하기 위해 특징점 매칭을 진행한다. 특징점 매칭은 세그먼트의 프레임 순서대로 진행하며 진행시 2개의 세그먼트가 필요하기 때문에 A와 B로 각 세그먼트에 대한 역할을 부여한다. 즉, 최초 특징점 매칭 시 프레임 상 첫번째 세그먼트가 A이고 두 번째 세그먼트가 B 역할을 수행한다. 매칭 인식률이 30% 이하로 떨어질 때까지 프레임 상 그 다음 세그먼트를 B로 설정한다. 인식률이 30% 이하로 내려갈 경우 해당 세그먼트를 A 역할로 재 배치하고 해당 세그먼트를 이미지 스티칭에 필요한 세그먼트로 분류한다.

분류된 약통 세그먼트 데이터셋은 연속적인 세그먼트들의 특징점들을 순차적으로 비교하여 이미지 스티칭을 진행하기 위해 슬라이딩 윈도우(Sliding Window) 기법으로 프로세스를 진행한다 (Fig. 3). 세그먼트 a 번째에게 A역할을, a+1 번째에게 B역할을 줌으로서 특징점을 비교하는 방식이다. 이미지 스티칭 과정에서 검출된 특징점에 RANSAC 거리행렬을 적용한 후 x축을 기점으로 오름차순으로 정렬한다. 좌측 이미지에 해당하는 A의 x 좌표 최댓값과 y 좌표 최솟값, 우측 이미지에 해당하는 B의 x좌표 최소값과 y좌표 최소값을 기준으로 이미지 스티칭을 진행한다.

특정 약통 식별 테스트 시, 이미지 스티칭으로 생성된 모델 이미지와 비교 이미지를 1픽셀 단위로 이동하면서 특정 약통에 대한 전체 프레임에서의 인식률을 테스트한다 (Fig 5). 수직 축 회전 영상에서 프레임을 기준으로 스티칭된 model data는 객체의 영역을 기준으로 이동한다.

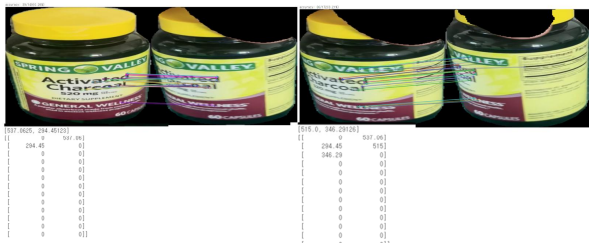


Fig. 3. Image Stitching Process

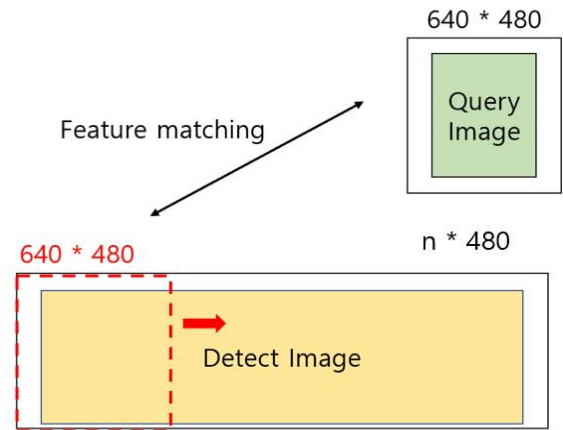


Fig. 4. pill bottle identification recognition test process



Fig. 5. four-frame stitching (left), six-frame stitching(right) feature matching result

IV. Result

아핀 분석 테스트(Fig. 6)를 통해 약통을 회전시키는 손의 움직임에 따라 값이 변화하는 것을 알 수 있었다. 회전값이 최대이면서 급격하게 상승한 값인 38.8과 146.3 지점에서는 약통을 회전시키기 위해 손을 떼고 다시 약통을 잡는 부분이다. 이 부분은 약통이 정지했다고 판단할 수 있다. 라플라시안(Laplacian) 분석 시 12초 동안 회전하는 영상에서 최솟값이 258을 기록했고 라플라시안 결과는 회전 속도에 영향을 받고 있기 때문에 천천히 회전하는 약통일수록 데이터 이미지 추출 과정에서 정확도를 높일 수 있다고 판단된다. 또한 특징점 매칭시 특징점 매칭률이 높은 점들만 추출하기 때문에 손가락이 약통의 표지에 크게 영향을 주지 않는다면 특징점 매칭시 정확도(인식률)에 치명적인 영향을 주지 않는다.(Fig 3)

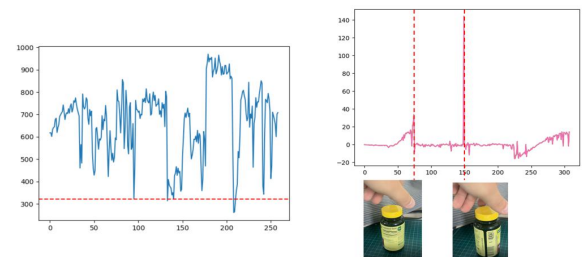


Fig. 6. Affine based image movement and rotation test (left), Laplacian test result (right)

Table 1. Recognition test result

		s4	s6	s8	s12	s6_orb	full_orb
Roatation Video	f1	0.272	0.272	0.311	0.300	0.217	0.267
	f2	0.304	0.304	0.304	0.304	0.272	0.264
	f3	0.315	0.197	0.315	0.315	0.324	0.248
	f4	0.147	0.147	0.327	0.305	0.088	0.14
	f5	0.328	0.328	0.328	0.328	0.095	0.125
	f6	0.095	0.105	0.168	0.246	0.010	0.109
YOLO v8 Segement	f1	1.00	1.0	1.0	1.0	0.253	0.912
	f2	0.920	1.0	0.925	1.0	0.301	0.917
	f3	0.20	1.0	0.325	1.0	0.132	0.534
	f4	1.00	1.0	1.0	1.0	0.114	0.335
	f5	0.532	1.0	0.532	1.0	0.132	0.15
	f6	0.460	1.0	0.55	1.0	0.08	0.08

f1~6 : video frame * n 번째
s4~s12 : segment image * n개로 stitching 된 model data image
s6_orb: segment 6개를 ORB로 합성한 model data image
full_orb: segment 이미지를 ORB로 분석 및 합성한 model data image

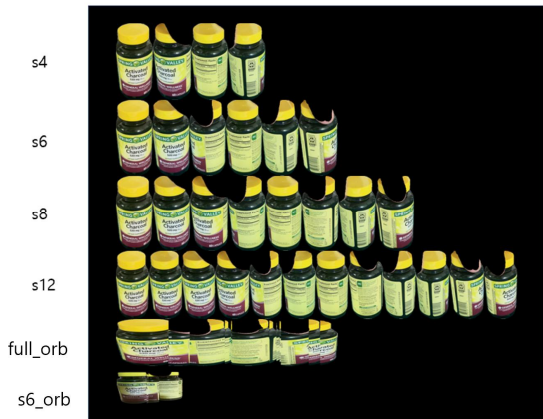


Fig. 7. Created model images

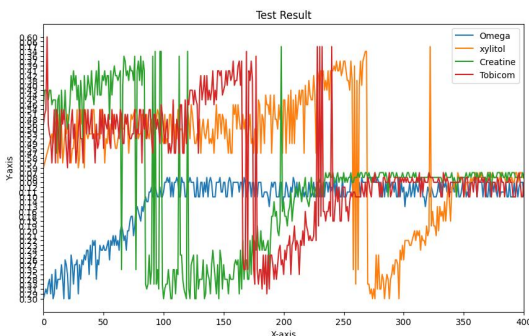
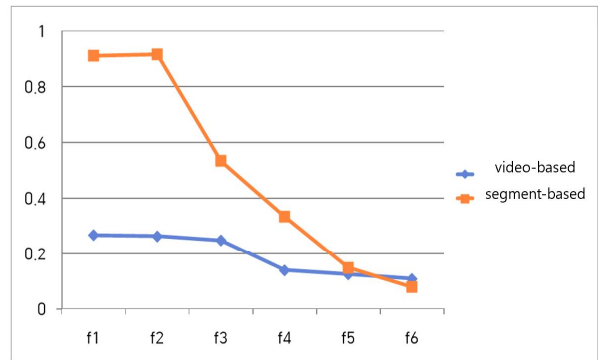


Fig. 8. Four object Accuracy test graph



	f1	f2	f3	f4	f5	f6
video-based	0.267	0.264	0.248	0.14	0.125	0.109
segment-based	0.912	0.917	0.534	0.335	0.15	0.08
편차	0.645	0.653	0.286	0.195	0.025	-0.029

Fig. 9. Comparison between video-based and segment-based query images

Fig 7와 Table 1은 각각 다른 방법으로 모델 이미지를 생성한 후 accuracy를 테스트한 결과이다. s4부터 s6은 프레임을 기준으로 추출된 약통 세그먼트 n개를 수평하게 붙혀 생선한 모델 이미지들이고, s6_orb는 프레임을 기준으로 추출된 약통 세그먼트 6개를 이미지 스티칭 과정 (Fig. 3)만을 거친 후 생성된 모델 이미지이며, full_orb는 3장에서 설계한 모든 과정을 거쳐 생성된 모델 이미지이다.

테스트의 목적 첫 번째의 프레임 개수별 인식률의 차이를 비교하는 s4~s12의 평균 인식률을 보았을 때 s4 0.2435, s6 0.2255, s8 0.2921, s12 0.2996으로 s12일 경우가 가장 인식률이 좋다고 판단할 수 있다 (Table 1). 하지만 정확도 향상을 위해 프레임 수를 늘리면 한 약통당 12장 이상의 이미지가 필요하기 때문에 메모리 공간에 제약이 생길 수 있다. 그렇기 때문에 적은 수의 이미지로 특정 약통을 식별할 수 있는 방법으로 스티칭된 이미지의 인식률도 비교한다. Table 2를 살펴보면 이미지에서 약통의 경계 영역도 accuracy에 영향을 미친다. Table 2의 video-based는 영상에서 약통이 인식되면 인식 영역의 x,y축을 기준으로 추출한 후 모델 이미지와 비교했을 때의 인식률이고 segment-based는 영상에서 약통이 인식되면 인식된 약통의 세그먼트를 추출한 후에 모델 이미지와 인식률을 비교한 것이다. 각 이미지들의 인식률의 편차를 살펴보면 최대 0.45까지 차이가 나는 것을 볼 수 있다. 이는 뒷 배경이 제거된 약통의 세그먼트를 추출한 후 특징 점 매칭을 진행했을 때 정확도가 향상됨을 의미한다. 하지만 동일하게 위에서 언급 바와 같이 약물 성분 부분의 노출이 증가할수록 인식률이 저하하기 시작하며 인식률의 편차도 0.025(절댓값)까지 떨어졌다.

테스트 시 f4, f5, f6에서 인식률이 점차 감소하는 것을 확인할 수 있다. 이는 회전하는 약통에서 특징을 검출할 때 특징점이 많이 분포하고 있는 약통의 정면부 노출이 감소할수록 특징점 분포가 적은 약통 성분표의 노출이 증가할수록 인식률이 떨어지고 있음을 볼 수 있다. 같은 현상으로 정면 프레임 노출이 줄어들고 문자들이 보이기 시작하는 f4의 테스트 결과를 살펴보았을 때, 특징점은 정상적으로 매칭되고 있으나 매칭되는 특징점이 충분하지 않으며 지역적으로 분포되어 있어 두 이미지의 매칭률이 0.147과 0.14로 현저히 떨어졌음을 볼 수 있다. 즉, 특정 약통을 식별의 정확도를 높이기 위해서는 특징점들이 약통의 전 범위에 많이 분포해야 한다는 것을 고려해야 한다.

Fig 9는 약통 4개를 정면을 기준으로 테스트했을 때 수직 회전 영상에서 프레임별 인식률 변화를 보여주는 그래프이다. 테스트 결과 약통의 수직축 회전 속도에 따라 y축 즉 정확도가 급격하게 떨어지는 부분에 차이가 있음을 알 수 있다. 정확도 고점의 범위는 0.52~0.59로 회전 속도와 상관없이 정확도가 50% 이상 기록이 될 경우에 안정적으로 영상에서 약통 정면이 인식되었다고 판단할 수 있다. 그래프의 데이터 추세는 정면 노출이 적어지는 부분을 인식했을 경우 예측률이 급격하게 떨어짐을 알 수 있고 다시 정면이 노출됨에 따라 예측률이 상승하고 있다. Fig 8에서 다른 추세를 보이는 약통은 빛이 반사되는 재질로 제작된 약 성분 표지 지에 의해 특징점이 텍스트나 도형 등 약통 커버가 아닌 빛의 반사 지점에 의해 특징이 추출되어 상대적으로 반사 지점이 적은 약통의 측면부에서 인식률이 높아지고 있음을 알 수 있다.

Fig 9과 Table 3은 테스트 목적 두 번째인, 스티칭된 약통의 인식률을 비교하는 그림과 표이며, 연구실에서 임의로 수집한 약통 10개의 ORB 스티칭된 이미지들을 통해 정확도를 비교한 결과이다. 테스트 결과 타겟 약통의 특징점 매칭의 최대 정확도는 0.792, 최소 0.345, 평균 0.593이다. 타겟 약통과 다른 약통과 대조를 진행했을 때의 정확도의 최댓값은 0.240, 최솟값 0.048, 평균 0.125로 타겟 약통이 아닌 약통을 매칭 했을 때 인식률을 0.3을 넘지 않고, 타겟 약통을 매칭했을 때 최소 인식률인 0.345를 보아 0.3 이상의 정확도를 기록하고 있음을 알 수 있다. 이를 통해 0.3을 임계값으로 설정하고 스티칭된 특정 약통 이미지와 타 약통의 특징점 매칭을 진행했을 때 0.3 이상의 정확도가 기록되면 특정 약통이 식별되었다는 결론을 도출할 수 있다. 혹은, 특징점 매칭 시 정확도가 가장 높은 약통이 타겟 약통이라고 지정할 시에도 같은 판단의 결과를 보여주고 있다. 즉 임계값 설정을 통해 약통을 식별한다면

빠르게 식별 결과를 필터링해나갈 수 있지만 적절한 임계값을 설정하지 않는다면 잘못된 판단을 할 수 있고 최대 정확도 검출 방법은 시간이 오래 소요되지만 정확한 결과를 얻을 수 있다.

Table 2. Accuracy comparison matrix between pill bottle segments and their model images

	A	B	C	D	E	F	G	H	I	J
A	0.370	0.141	0.127	0.148	0.138	0.122	0.109	0.121	0.122	0.117
B	0.108	0.376	0.112	0.176	0.138	0.116	0.127	0.105	0.106	0.111
C	0.115	0.171	0.605	0.171	0.154	0.240	0.140	0.124	0.133	0.101
D	0.112	0.152	0.126	0.345	0.129	0.109	0.156	0.106	0.127	0.048
E	0.111	0.130	0.120	0.130	0.547	0.123	0.113	0.107	0.112	0.092
F	0.133	0.162	0.123	0.142	0.127	0.766	0.141	0.108	0.115	0.096
G	0.109	0.140	0.147	0.156	0.148	0.135	0.792	0.111	0.117	0.115
H	0.111	0.144	0.121	0.147	0.128	0.106	0.115	0.610	0.136	0.101
I	0.103	0.122	0.114	0.133	0.124	0.129	0.103	0.107	0.762	0.093
J	0.110	0.130	0.117	0.136	0.125	0.115	0.123	0.097	0.119	0.760

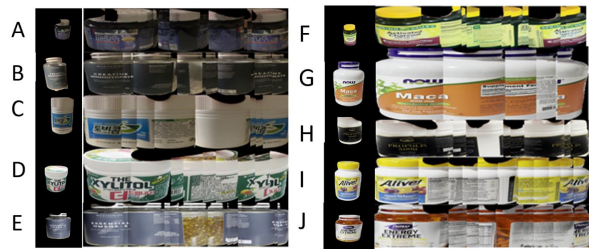


Fig. 10. Pill bottle stitching result

V. Conclusions

본 연구를 통해 YOLOv8의 세그먼트 인식 기술과 특징점 매칭의 ORB 알고리즘을 결합하는 시스템으로 약통을 인식 및 식별하는 방법을 제안하였다. 이를 통해 자동으로 영상에서 객체를 인식하고 객체에 대한 다각도 이미지가 통합된 단일 이미지를 특정 약통에 대한 식별에 활용할 수 있다는 결과를 얻었다. 영상에서 수집한 세그먼트들을 ORB를 통해 스티칭하여 개별 약통에 대한 모델 이미지를 생성하였으며 정면 인식률 0.912, 측면 인식률 0.355으로 특징점 추출에 제약이 있는 약물설명 부분(후면)의 노출도를 낮추면 특정 약통에 대한 식별이 가능함을 보였다. 하지만 약통 전면에 대한 인식률 평균을 높이기 위해서는 시차 연산과 회전 속도를 고려해야 할 필요가 있다. 모델 이미지가 스티칭 되는 과정에서 표본 선정과 전처리 단계에서 이를 개선한다면 봉안된 시스템이 설계될 수 있을 것이라고 기대한다.

한편으로 세그먼트의 특징점 매칭 인식률의 임계값으로 분류한 데이터 집합은 영상에서의 약통과 특징점 매칭 분

석을 통해 해당 시스템이 객체의 다각도의 데이터를 수집할 수 있음을 보여준다. 데이터 집합 군이 많을수록 인식을 증가하기 때문에 많은 메모리 공간이 필요하게 되며 모든 이미지 데이터를 순회하면 런타임이 증가하게 된다. 데이터 집합 군을 형성할 때 객체의 각도와 조도 등을 고려하여 이미지 자료수집 수 있다면 불필요한 이미지를 필터링하여 메모리를 효율적으로 관리할 수 있을 것으로 예측된다.

향후 위와 같은 본 식별 시스템의 제약조건인 시차 연산 제약조건을 개선하여 특정 객체 식별 정확도를 향상시킬 수 있는 시스템을 설계할 계획이며 원통형 약통뿐만 아니라 다각형 약통(약상자)의 특정 객체 식별을 위해 해당 시스템을 적용하여 시스템을 개선해나갈 계획이다.

REFERENCES

- [1] Joseph Redmon, and Sentosh Divvala, "You Only Look Once: Unified, Real-Time Object Detection", arXiv:1506.02640, 9 May 2016
- [2] Juan Terven, Diana Cordova-Esparza, "YA Comprehensive Review of YOLO: From YOLOv1 and Beyond", arXiv:2304.00501, 2 Apr 2023
- [3] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," 2011 International Conference on Computer Vision, Barcelona, Spain, 2011, pp. 2564-2571, doi: 10.1109/ICCV.2011.6126544.
- [4] Calonder, M., Lepetit, V., Strecha, C., Fua, P. (2010). BRIEF: Binary Robust Independent Elementary Features. In: Daniilidis, K., Maragos, P., Paragios, N. (eds) Computer Vision - ECCV 2010. ECCV 2010.
- [5] Ultralytics, Ultralytics YOLOv8 Docs <https://docs.ultralytics.com/>
- [6] Q. Deng, Y. Luo, and J. Ge, "Dual threshold based unsupervised face image clustering," Proceedings of the 2nd International Conference on Industrial Mechatronics and Automation, pp. 436-439, 2010.
- [7] Kyuho Sim, & Guihyung Lee (2021). Neural-network-based Object Recognition and RANSAC-based Three-dimensional Posture Estimation for Solving the Pickup Problem of Occluded Parts. Journal of the Korean Society of Manufacturing Technology Engineers, 30(5), 372-381, 10.7735/ksmte.2021.30.5.372
- [8] JO, Ikhyeon, SA, Jaemin, KIM, Yongyu, & NAM, Baek (2022). Improve Ultra-low Latency Vehicle Identification and Traffic Precision Based Yolo-v4. Journal of Korean Society of Transportation, 40(6), 908-920, 10.7470/jkst.2022.40.6.908
- [9] Joo, Seong Moon, and Lee, Chil woo. "3D Object Recognition Using Appearance Model Space of Feature Point." KIPS Transactions on Software and Data Engineering 3, no. 2 (February 28, 2014): 93-100. doi:10.3745/KTSDE.2014.3.2.93.
- [10] L. Magalhães, B. Ribeiro, N. Alves and M. Guevara, "A three-staged approach to medicine box recognition," 2017 24^o Encontro Português de Computação Gráfica e Interação (EPCGI), Guimaraes, Portugal, 2017, pp. 1-7, doi: 10.1109/EPCGI.2017.8124317.
- [11] D. Hu and M. Pan, "Research on a Algorithm of Medicine Box Tracing Code Location," 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2023, pp. 124-128, doi: 10.1109/ICCECE58074.2023.10135290
- [12] Edward Rosten, Reid Porter, and Tom Drummond, "Faster and better: a machine learning approach to corner detection" in IEEE Trans. Pattern Analysis and Machine Intelligence, 2010, vol 32, pp. 105-119.
- [13] Jianfei Zhang, Sai Ke, "Improved YOLOX Fire Scenario Detection Method", Hindawi Wireless Communications and Mobile Computing, Volume 2022, Article ID 9666265, 8 pages
- [14] L. Magalhães, B. Ribeiro, N. Alves and M. Guevara, "A three-staged approach to medicine box recognition," 2017 24^o Encontro Português de Computação Gráfica e Interação (EPCGI), Guimaraes, Portugal, 2017, pp. 1-7, doi: 10.1109/EPCGI.2017.8124317.

Authors



Dae-Hyun Kim received the A.S. degrees in Department of Computer Science from Inha Technical College, Korea, in 2023 and is currently a B.S. in Department of Computer Science and Engineering at Inha Technical

College, Incheon Korea. His research interests include artificial intelligence, multimodal system, and computer vision.



Hyo Hyun Choi received his B.S., M.S., and Ph.D. degrees in computer science and engineering from Sogang University, Seoul, Korea in 1994, 1996, and 2005, respectively. He joined Telecommunication R&D Center of

Samsung Electronics as a senior engineer in 2005, where he was in charge of the research on the wireless sensor network. In 2009, he joined the faculty of the Department of Computer Science, Inha Technical College, Incheon, Korea, as a professor. His research interests are artificial intelligence, wireless sensor network, advanced air mobility, ubiquitous computing, and blockchain technology.