

딥러닝을 활용한 잔량 분포에 따른 도시락 부피 측정 연구

Study on Measuring Lunchbox Volume Based on Residual Distribution Using Deep Learning

박승빈¹, 김광기^{2*}

Seung-Been Park¹, Kwang-Gi Kim^{2*}

요약

이 연구는 노인의 영양 섭취를 효과적으로 모니터링할 수 있는 시스템을 개발하는 데 중점을 두었으며, 이는 Flutter 기반의 사용자 인터페이스와 U-Net 모델을 활용한 이미지 처리 백엔드로 구성된다. 사용자는 도시락의 이미지를 앱에 업로드하면, 시스템은 U-Net 모델을 통해 이미지 내의 잔량을 자동으로 분석하고 결과를 제공한다. 시스템 검증에 위해 잔량을 25%, 50%, 75%, 100%로 분류하여 각각에서의 분할 성능을 비교하였고, 이 과정에서 U-Net 모델은 매우 정밀한 이미지 분할 성능을 보이며 다양한 실험을 통해 높은 정확도를 기록하였다. 본 시스템은 노인이 혼자 식사할 때 발생할 수 있는 영양 부족 문제를 해결하기 위해 설계되었으며, 실시간으로 잔량을 모니터링함으로써 전문가들이 보다 적절한 영양 관리를 할 수 있도록 돕는다. 이러한 연구 결과는 취약 계층의 영양 상태를 체계적으로 모니터링하고 개선하는 데 있어 중요한 도구로 활용될 가능성을 보여준다. 이러한 기술 기반의 접근은 공공 보건 영역에서 노인의 영양 모니터링을 자동화하고, 실시간 개입을 가능하게 하는 중요한 전환점을 마련할 수 있다.

핵심어 : 도시락 모니터링, 영양섭취, U-Net, 이미지 분할

Abstract

In this paper, we focus on developing a system that can effectively monitor the nutritional intake of the elderly. The system consists of a Flutter-based user interface and an image processing backend using a U-Net model. When a user uploads an image of a lunchbox to the application, the system automatically analyzes the remaining amount in the image through the U-Net model and provides the result. For system validation, the remaining amounts were classified into 25%, 50%, 75%, and 100%, and the segmentation performance at each level was compared. In this process, the U-Net model showed very precise image segmentation performance and recorded high accuracy through various experiments. This system was designed to solve the problem of nutritional deficiency that may occur when elderly individuals eat alone. By monitoring the remaining amount in real time, it helps experts provide more appropriate nutritional

1 Department Biomedical Engineering, Gachon University, Seoungnam, Korea [Graduate Student]

e-mail: jh86237@gachon.ac.kr

2 Department Biomedical Engineering, Gachon University, Seoungnam, Korea [Professor]

e-mail: kimkg@gachon.ac.kr (Corresponding author)

*이 논문은 2024년도 길의료재단의 재원으로 [형광발현 유도하 수술용 로봇 촬영 시스템 탑재를 위한 인공지능 기반 실시간 형광발현 림프노드 검출 모델 개발]의 지원을 받아 수행된 연구임(FRD2024-13).

Received(January 21, 2025), Review Result(1st: February 18, 2025), Accepted(April 11, 2025), Published(April 30, 2025)



© 2025 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

management. These research results demonstrate the potential of this system as an important tool for systematically monitoring and improving the nutritional status of vulnerable groups. This technology-based approach can mark a significant turning point in the public health sector by automating nutrition monitoring for the elderly and enabling real-time intervention.

Keyword : Lunchbox Monitoring, Nutritional Intake, U-Net, Image Segmentation

1. 서론

취약계층의 영양 결핍은 약 44.7%의 노인에게서 나타나는 심각한 공중보건 문제로, 이는 건강에 치명적인 영향을 미칠 수 있다 [1]. 체질량지수(Body Mass Index, BMI)나 혈청 알부민 수치만으로는 영양 결핍을 정확히 평가하기 어려운 경우가 많아, 이를 보완할 효율적인 모니터링 방법이 필요하다 [2].

전통적으로 취약계층의 도시락 잔량을 확인하는 방법은 주로 방문 조사나 자가 보고를 하는 방식으로 이루어졌다 [3]. 2022년, Jelna Lubura 등은 컨볼루션 신경망(Convolution Neural Network, CNN) 모델을 활용해 20세에서 30세 사이의 사람들을 대상으로 하루 식사 후 음식 잔량을 평가하였고, 98% 이상의 높은 정확도를 보였다 [4]. 2023년, Yuita Arum Sari 등은 CNN을 사용하여 픽셀 기반 AFLE(Adaptive Food Leftover Estimation) 및 회귀 접근 방식을 통해 음식의 잔량을 추정하였다. 드롭아웃 0.3과 이미지 증강(Image augmentation), ReLU 활성화 함수를 사용하였을 때 92.5%의 높은 정확도를 보였다 [5].

2016년 Weishan Zhang 등은 5개의 층으로 구성된 CNN을 사용하여 음식 이미지 인식 시스템을 설계하였으며, UEC-FOOD100 데이터 세트(약 15,000개의 이미지)와 자체 구축한 과일 데이터 세트(약 40,000개의 이미지)를 활용하여 실험을 진행했다. 해당 연구에서는 다중 음식 데이터 세트에서 60.9%의 정확도를, 과일 데이터 세트에서 80.8%의 정확도를 달성하여 CNN의 가능성을 보여주었다 [6].

2015년 Marios M.Anthimopoulos의 연구는 선형 서포트 벡터 머신(Support Vector Machine, SVM)을 사용해 약 78%의 정확도로 음식을 분류했다 [7]. 2015년 Ye He, Chang Xu 등은 Local Variation 기법을 사용해 식품 분류를 하였다. 해당 논문에서는 정형 식품에 대해 3D 재구성 방법을 사용하고, 비정형 식품이나 형태 변동이 큰 식품에 대해서는 면적 기반 무게 추정 방법을 사용하는 방법을 제안하였다 [8].

U-Net 기반 연구도 점차 주목받고 있다. 예를 들어, 2021년 Kyamelia Roy 등은 딥러닝 아키텍처를 활용하여 사과의 껍질에 존재하는 결함을 기반으로 썩은 사과와 신선한 사과를 감지하는 방법을 제안했다. En-Unet 모델이 97.46%의 훈련 정확도와 0.866의 평균 Intersection over Union(IOU) 점수를 달성하여 실시간 분할 및 감지가 가능함을 보여주었다 [9]. 2024년 Xiongwei Wu 등은 OVFoodSeg라는 모델을 제안하여, 시각-언어 모델(VLMs)을 활용해 텍스트 임베딩을 이미지 맥락과

결합하고, 새로운 재료에도 유연하게 대응할 수 있는 열린 어휘(Open-vocabulary) 방식을 채택한 음식 이미지 분할을 수행한 바 있다. 이 접근법은 FoodSeg103 데이터 세트에서 평균 교차 비율(mIoU)을 4.9% 향상하며 성능을 크게 개선한 사례도 있다 [10].

칼라 영상 분할 사례로는 2020년 Zhuokun Pan 등의 무계획 도시 정착지에서 개별 건물의 분할과 분류를 위해 U-Net을 적용한 연구가 있다. 연구 결과, 건물 분할의 전체 정확도는 86% 이상, 분류 정확도는 83% 이상을 기록하며, 기존의 랜덤 포레스트 및 객체 기반 이미지 분석 방법보다 성능이 우수함을 보여주었다 [11].

CNN에 비해 U-Net을 활용하였을 때 음식 외 요소를 제거하기 용이하고, 보다 정밀한 이미지 분할이 가능하다. 그러나 U-Net 모델을 사용하여 도시락의 잔량을 분할하고 애플리케이션을 구현한 연구 사례는 많지 않다. 본 논문에서는 U-Net 딥러닝 모델을 사용하여 도시락 안에 있는 밥의 잔량을 측정하고 각각 25%, 50%, 75%, 100% 분량의 잔량이 남았을 때의 결과를 분석하였다.

2. 방법

본 연구에서는 취약계층의 영양 섭취를 모니터링하기 위해 플러터(Flutter)를 사용한 프론트엔드(Frontend)와 U-Net 기반의 백엔드(Backend) 이미지 처리 기술을 결합한 애플리케이션을 개발하였다. 이를 통해 도시락 사진을 분석하여 남은 음식의 양을 측정하고, 일, 주, 월별 영양 섭취 현황을 그래프로 시각화하여 취약계층의 영양 상태를 효과적으로 파악할 수 있도록 하였다.

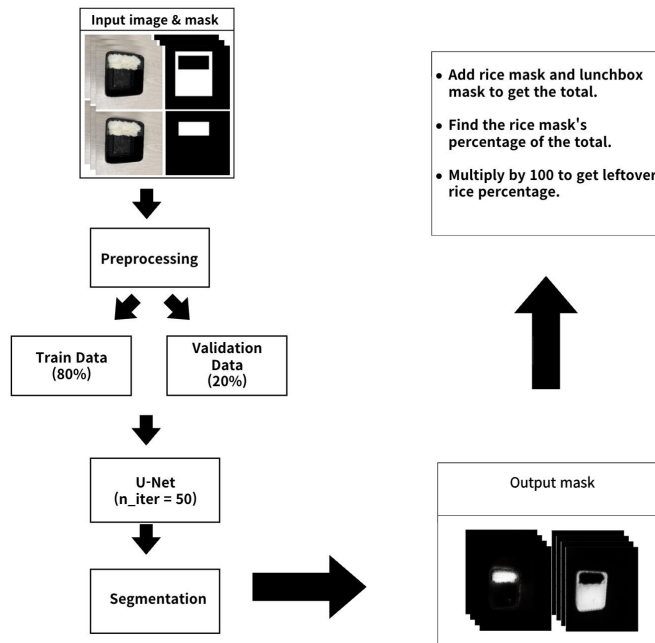
사용자가 애플리케이션을 통해 도시락 사진을 촬영하면, 해당 이미지는 서버로 전송된다. 서버에서는 전송된 이미지를 U-Net 모델을 사용하여 밥과 도시락을 분할하고, 각 영역의 픽셀 수를 예측하여 남은 음식의 양을 측정한다. 최종적으로, 서버는 예측된 결과를 실제 크기에 맞게 정규화하여 애플리케이션 내에서 시각적으로 사용자에게 제공한다. [그림 1]은 도시락 잔량 분석을 위한 딥러닝 기반 분할 모델의 작업 흐름도이다.

애플리케이션의 프론트엔드는 플러터를 사용하여 개발되었다. 플러터는 크로스 플랫폼 프레임워크로, Android와 IOS에서 모두 동작할 수 있는 애플리케이션을 개발할 수 있다. 주요 기능은 사용자가 도시락의 사진을 촬영하고, 촬영된 이미지를 서버로 업로드하는 것이다.

애플리케이션에서 사용한 주요 패키지는 image_picker와 http이다. image_picker를 사용하여 사용자가 기기의 카메라를 이용해 이미지를 촬영할 수 있도록 하였으며, 촬영된 이미지는 로컬 스토리지에 저장된다. 이후, HTTP POST 요청을 통해 이미지를 서버로 전송한다. 사용자는 업로드된 이미지를 확인할 수 있으며, 서버로부터 응답을 받아 업로드 성공 여부를 확인할 수 있다.

백엔드는 Flask 프레임워크와 Flask-CORS를 사용하여 구축되었으며, 플러터 애플리케이션에서 촬영된 도시락 이미지를 처리한다. 사용자가 촬영한 이미지는 HTTP POST 요청을 통해 서버로 전

송되며, 서버는 전송된 이미지를 지정된 디렉토리에 저장하여 관리한다. 이미지가 서버에 업로드되면, 사전 학습된 U-Net 모델이 로드되고, 이미지를 분석하는 과정을 진행한다.



[그림 1] 도시락 잔량 분석을 위한 딥러닝 기반 분할 모델의 작업 흐름도

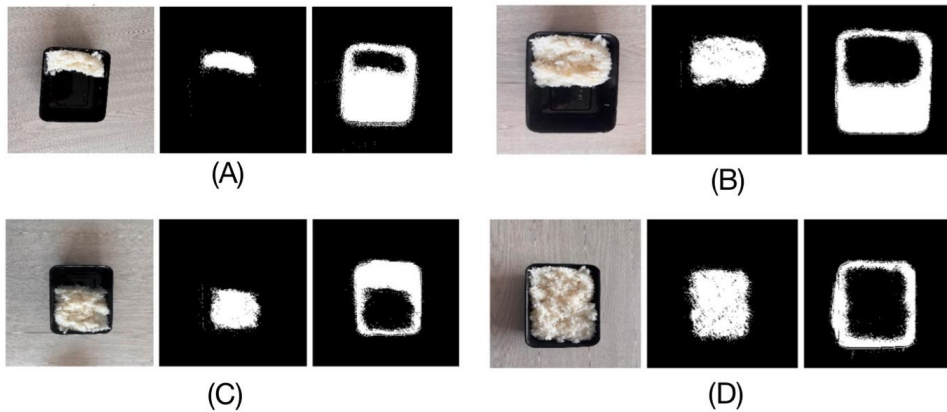
[Fig. 1] Workflow of the Deep Learning-Based Segmentation Model for Lunchbox Leftover Analysis

이미지는 텐서 형식으로 변경되어 U-Net 모델에 전달된다. 이 모델은 입력된 도시락 이미지에서 밥과 도시락의 영역을 각각 분리하는 작업을 수행한다. 모델이 분리한 각 영역의 픽셀 단위 면적을 계산하여, 밥이 도시락 전체 면적에서 차지하는 비율을 산출한다. 이 비율을 통해 밥의 잔량을 파악할 수 있으며, 계산된 밥의 면적 비율은 백분율로 표현된다. 서버는 결과를 JSON 형식으로 반환하며, 여기에는 650ml 용량의 용기에 맞춰 정규화된 밥의 양에 대한 정보도 포함되었다.

[그림 2]는 도시락에서 밥 비율에 대한 모델의 예측을 시각화한 것이다.

본 연구에서는 도시락과 밥의 영역을 정확하게 분할하기 위해 U-Net 구조를 기반으로 한 심층 신경망을 사용하였다 [12]. U-Net은 의료 영상 처리 분야에서 널리 사용되는 구조로, 복잡한 물체의 경계를 효과적으로 학습할 수 있는 장점을 가지고 있다. 본 연구에서는 이 구조를 응용하여 도시락과 밥의 픽셀 단위 분할을 수행하였다. 모델 학습을 위해 도시락 이미지와 해당 이미지에 대한 도시락 및 밥 영역 마스크를 각각 준비하였다. 데이터는 다양한 각도와 조명 조건을 반영하도록 수집하였으며, 데이터셋의 다양성을 높이기 위해 데이터 증강 기법을 적용하였다. 증강 기법으로는 수평 및 수직 반전, 회전, 이동, 확대 및 축소 변환 등을 사용하여 학습 데이터의 양을 증가

시키고 모델의 일반화 성능을 향상시키고자 하였다.



[그림 2] 도시락에서 밥 비율에 대한 모델의 예측 시각화: (A) 25%, (B) 50%, (C) 75%, (D) 100%

[Fig. 2] Visualization of the model's predictions for rice proportions in a lunchbox set to (A) 25%, (B) 50%, (C) 75%, and (D) 100%

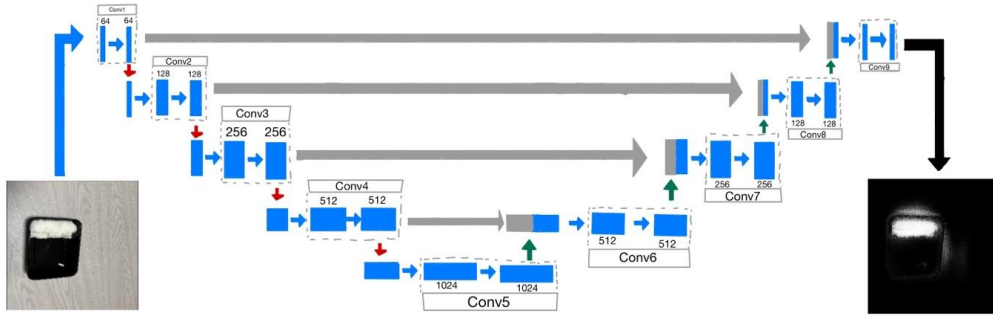
각 이미지는 256x256 크기로 조정되었으며, 마스크 이미지의 경우, 특정 레이블(밥과 도시락)을 가진 픽셀을 1로 설정하고, 나머지 영역을 0으로 설정하여 이진 마스크를 생성하였다. 이를 통해 모델이 각 픽셀의 클래스(도시락, 밥)를 정확하게 학습할 수 있도록 하였다. U-Net 모델은 대칭적인 인코딩-디코딩 구조를 가지며, 인코더 부분에서는 컨볼루션과 풀링 레이어를 사용하여 이미지의 특성을 추출하고, 디코더 부분에서는 업샘플링과 컨볼루션 레이어를 통해 이미지를 복원하면서 세그멘테이션을 수행한다. 인코더와 디코더 사이의 스킵 연결(Skip connection)은 낮은 레벨에서 학습된 정보를 디코더로 전달하여 더욱 정확한 분할이 가능하도록 한다.

본 연구에서는 다음과 같은 U-Net 구조를 사용하였다:

인코더: 4개의 컨볼루션 블록(각 블록은 두 개의 3x3 컨볼루션 레이어와 ReLU 활성화 함수로 구성)과 풀링 레이어로 구성한다. 각 블록을 지나면서 채널 수는 두 배로 증가하며, 공간 해상도는 절반으로 감소한다.

디코더: 4개의 업샘플링 블록(각 블록은 업샘플링 레이어, 두 개의 3x3 컨볼루션 레이어, 그리고 ReLU 활성화 함수로 구성)으로 구성한다. 업샘플링 후에는 해당 단계에서 인코더 블록의 출력을 컨캐티네이션(Concatenation)하여 세그멘테이션(Segmentation) 정확도를 높이고자 하였다.

출력 레이어: 마지막 컨볼루션 레이어는 1x1 커널을 사용하여 각 픽셀에 대해 두 개의 클래스(밥, 도시락)에 대한 확률을 출력하며, 시그모이드 활성화 함수를 사용하여 최종 예측 값(분할영상)을 얻도록 하였다. [그림 3]은 U-Net 모델의 구조에 대한 설명이다.



[그림 3] 밥과 도시락 분할을 위한 U-Net 기반 모델의 구조

[Fig. 3] Structure of the U-Net-Based Model Used for Rice and Lunchbox Segmentation

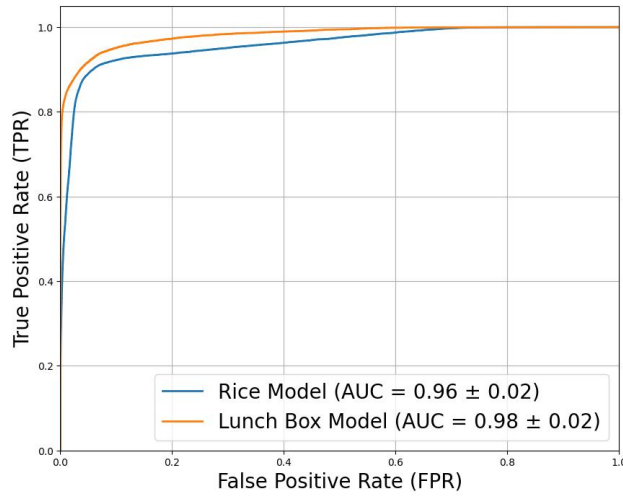
모델 학습은 Adam 최적화 알고리즘을 사용하여 진행되었으며, 학습률은 $1e-4$ 로 설정하였다. 손실 함수로는 이진 크로스 엔트로피(Binary Cross-Entropy, BCE)를 사용하였으며, 정확도 평가를 위해 Dice 계수를 활용하였다. Dice 계수는 두 세트 간의 겹치는 정도를 측정하여, 모델이 예측한 분할 영역이 실제 영역과 얼마나 일치하는지를 정량적으로 평가하는 메트릭이다.

학습 과정에서는 밥 영역과 도시락 영역 각각에 대해 BCE 손실 함수를 독립적으로 적용하여 각 클래스별 손실을 계산하였다. 이를 기반으로 모델의 가중치를 업데이트하며 학습을 최적화하였다. 총 50번의 반복(Iteration)을 통해 모델 학습을 하였으며, 매 반복마다 학습 손실 및 정확도를 추적하여 학습 과정의 성능을 평가하였다.

3. 결과

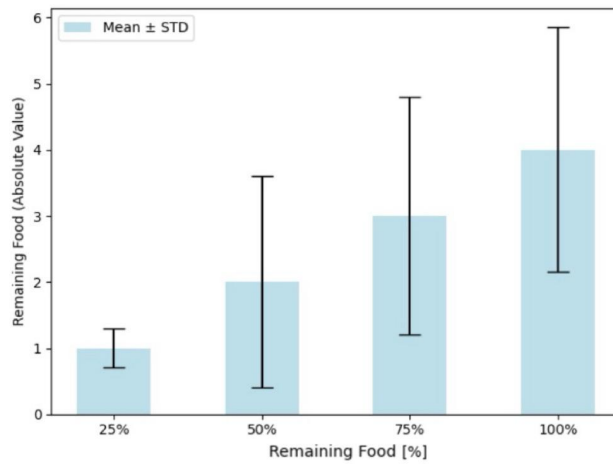
모델 학습 후, 검증 데이터셋을 활용하여 성능을 평가하였다. 검증 데이터셋에서 모델의 예측 결과를 실제 값과 비교하여 Dice 계수, 정확도(Accuracy), 민감도(Sensitivity), 특이도(Specificity)와 같은 지표를 계산하였다. 이를 통해 모델이 이미지 분할 문제를 얼마나 효과적으로 수행했는지를 확인할 수 있었다.

또한, 검증 세트에 대한 정확도를 밥 비율에 따라 분석한 결과, 밥 비율이 낮을 때는 정확도가 높았으며, 반면 밥 비율이 증가할수록 정확도가 상대적으로 낮아지는 경향을 보였다. 이와 함께, 밥 비율이 증가할수록 표준 편차(STD)도 증가하는 양상을 보여, 밥과 도시락 간 경계가 모호해지면서 모델의 예측 성능이 더 큰 변동성을 가지게 되었음을 시사한다. 이는 ROI(Region of Interest) 과정에서 밥 비율이 많아질수록 도시락과 밥의 마스크가 겹치는 경우가 증가하기 때문으로 추측된다. [그림 4]는 도시락과 밥 분할 모델에 대한 ROC 커브를 나타낸 것이고, [그림 5]는 밥 비율에 따른 모델의 표준 편차를 나타낸 것이다. 또한 [표 1]은 U-Net 및 Efficient U-Net 모델을 사용한 밥과 도시락 분할 이미지의 성능 지표를 나타낸다.



[그림 4] 도시락과 밥 분할 모델에 대한 ROC 커브

[Fig. 4] ROC Curve for the Rice and Lunchbox Segmentation Model



[그림 5] 밥 비율에 따른 분할 모델의 표준 편차

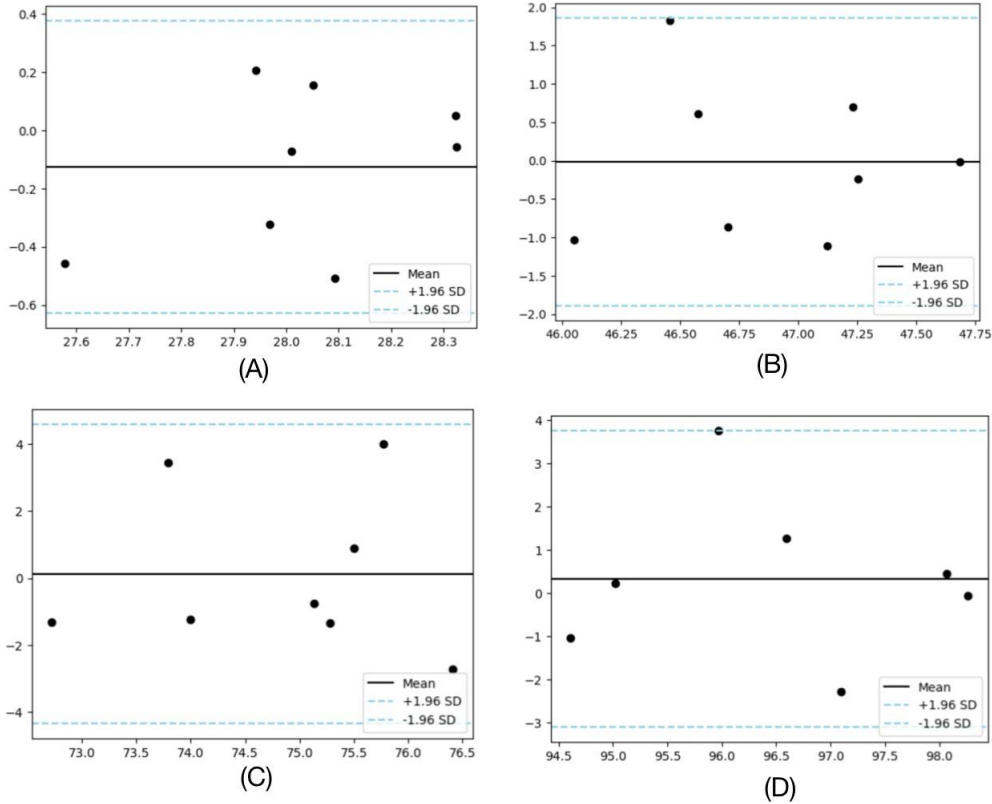
[Fig. 5] Standard Deviation of the Model According to Rice Proportion

[표 1] U-Net 및 Efficient U-Net 모델을 사용한 밥과 도시락 분할 이미지의 성능 지표

[Table 1] Performance Metrics for Rice and Lunchbox Segmentation Images using U-Net and Efficient U-Net Models

Model	dataset	sensitivity(%)	Accuracy (%)	Dice Coefficient(%)
U-Net	Rice	0.86 ± 0.04	0.98 ± 0.02	0.88 ± 0.05
Efficient-Net	Rice	0.98 ± 0.02	0.78 ± 0.01	0.77 ± 0.02
U-Net	Lunch-box	0.84 ± 0.03	0.97 ± 0.01	0.86 ± 0.04
Efficient-Net	Lunch-box	0.83 ± 0.03	0.72 ± 0.02	0.87 ± 0.02

개발된 시스템의 정확성을 검증하기 위해, 사전에 수집된 밥 잔량 데이터(25%, 50%, 75%, 100%의 밥 잔량 사진)를 사용하여 시스템의 일관성을 평가하였다. 각 잔량 상태에서 동일한 각도로 8번씩 촬영한 이미지를 활용하여, 모델의 예측값과 실제 측정값 간의 오차를 분석하였다.



[그림 6] 도시락에서 25%, 50%, 75%, 100%의 잔량 비율에 대해 동일한 모델을 여덟 번 측정한 결과. x축은 측정된 비율의 범위를 나타내고, y축은 각 측정값의 평균으로부터의 편차를 나타냄. (A) 25% 잔량 비율에서의 블랜드-알트만 그래프, (B) 50% 잔량 비율에서의 블랜드-알트만 그래프, (C) 75% 잔량 비율에서의 블랜드-알트만 그래프, (D) 100% 잔량 비율에서의 블랜드-알트만 그래프

[Fig. 6] Results of measuring the same model eight times at each leftover ratio of 25%, 50%, 75%, and 100% in a lunchbox. The x-axis represents the range of measured ratios, and the y-axis represents the deviation from the mean of each measurement. (A) Bland-Altman graph at 25% leftover ratio, (B) Bland-Altman graph at 50% leftover ratio, (C) Bland-Altman graph at 75% leftover ratio, (D) Bland-Altman graph at 100% leftover ratio

본 연구에서는 블랜드-알트만(Bland-Altman) 분석을 통해 알고리즘의 일관성을 평가하였다. 데이터는 25%, 50%, 75%, 100%의 밥 잔량을 동일한 촬영 각도에서 각각 8번씩 촬영한 후, 이 데이터를 바탕으로 모델의 예측값과 실제 측정값 간의 편차를 분석하였다. 블랜드-알트만 그림에서 X축

은 모델이 예측한 평균 잔량 비율을 나타내고, Y축은 예측값과 실제 측정값 간의 차이를 나타낸다. 그래프에서의 중앙선은 모델이 예측한 잔량의 평균을 의미하며, 이 선을 기준으로 위쪽에 위치한 값은 모델이 실제보다 더 많이 예측한 것이고, 아래쪽에 위치한 값은 더 적게 예측한 경우를 의미한다.

그래프에서 드러나는 몇몇 이상치는 촬영 각도가 미세하게 흔들리거나, 조명 등 외부 요인에 의한 변동으로 인해 발생할 수 있으며, 이러한 편차는 플롯에서 시각적으로 확인할 수 있다.

블랜드-알트만 분석 결과, 대부분의 데이터가 신뢰구간(± 1.96 표준 편차) 내에 위치하였으며, 이는 알고리즘이 일관성 있게 예측을 수행하고 있음을 나타낸다. 신뢰구간 내에 데이터가 밀집해 있다는 것은 동일한 조건에서 모델이 일관된 성능을 유지하고 있음을 시사하며, 시스템의 정확성과 신뢰성이 높다는 것을 보여준다. [그림 6]은 도시락 잔량 비율에 대한 블랜드-알트만표를 나타낸다.

4. 결론

취약 계층의 도시락 잔량을 확인하는 전통적인 방법은 주로 방문 조사나 자가 보고 방식에 의존해 왔으나, 이러한 방식은 인력 의존도가 높고 비용이 많이 들며, 특히 지역적으로 고립된 노인들을 자주 확인하기 어렵다는 한계가 있다. 이러한 문제를 해결하기 위해 최근에는 머신러닝 및 딥러닝 기반의 인공지능 기술을 활용하여 도시락 잔량을 확인하는 연구가 활발히 진행되고 있다.

이번 연구에서는 취약 계층의 영양 섭취를 모니터링하기 위해 도시락 잔량을 측정하는 애플리케이션을 개발하고, OTSU Algorithm(OTSU) 등 다양한 이미지 처리 기법을 비교한 결과 U-Net 모델이 가장 적합한 방법임을 확인하였다. 그러나 연구 과정에서 몇 가지 한계점이 있어, 향후 연구에서는 이를 개선할 필요가 있다.

먼저, 조명에 대한 한계가 있었다. OTSU 방법을 적용한 초기 연구에서는 조명 조건에 매우 민감한 반응을 보였는데, 특히 실제 환경에서 조명의 밝기에 따라 이미지의 명암이 크게 달라지면서 밥과 도시락 영역의 정확한 분할이 어려웠다. 이러한 이유로 OTSU 방법은 실제 환경에서의 적용이 어렵다는 결론을 내렸다. 반면 U-Net 모델은 다양한 조명 조건에서도 비교적 안정적인 성능을 보였지만, 극단적인 조명 조건에서는 분할 정확도가 떨어질 수 있어 이를 보완하기 위한 조명 보정 기법의 추가적인 연구가 필요하다.

또한 깊이 측정의 한계도 있었다. 본 연구에서 사용한 이미지 처리 기법은 2차원 이미지에서 밥의 면적을 기반으로 잔량을 계산하였으나, 밥의 깊이를 고려하지 않았기 때문에 실제 밥의 양과는 오차가 발생할 가능성이 있다. 특히, 밥이 용기 안에서 쌓여있는 경우 2차원 면적만으로는 정확한 양을 추정하기 어려울 수 있다. 따라서 향후 연구에서는 깊이를 측정할 수 있는 3D 이미지 처리 기법 또는 깊이(Depth) 센서를 활용한 방법을 고려해야 한다.

향후 연구에서는 U-Net 모델의 변형이나 개선된 버전을 도입하는 것이 중요할 것이다. 특히, Attention U-Net을 도입하면 이미지 내에서 중요한 부분에 더 많은 가중치를 부여하여 중요한 세부 사항을 강조할 수 있다. 이러한 접근 방식은 조명 조건의 변화나 이미지의 복잡성에 불구하고, 더 정교한 분할 결과를 도출할 수 있게 해준다. Attention 메커니즘을 통해 모델이 중요하지 않은 배경 요소를 무시하고, 식사 잔량과 같은 핵심 요소에 집중하게 함으로써, 분할의 정확도를 높일 수 있다 [13][14].

또한, Res-UNet과 같은 모델은 잔차 연결(Residual connection)을 사용하여 네트워크의 학습 과정에서의 안정성을 높인다. 잔차 연결은 깊은 신경망에서 발생할 수 있는 기울기 소실 문제를 완화하며, 더 깊은 네트워크 구조를 효과적으로 훈련시킬 수 있게 해준다. 이를 통해 Res-UNet은 다양한 스케일에서의 특징을 더 잘 포착하고, 복잡한 이미지에서도 분할 성능을 유지할 수 있을 것이다 [15].

EfficientNet을 활용한 실험에서는 기대보다 저조한 성능을 보였다. EfficientNet은 주로 이미지 분류에 최적화된 모델로, 세분화 작업에서는 U-Net처럼 픽셀 단위의 세밀한 분할에 적합하지 않을 수 있다. 이는 EfficientNet이 복잡한 패턴을 처리하는 데 강점을 가지지만, 잔량 측정과 같은 단순한 경계 분할 작업에서는 오히려 과적합 되었을 가능성이 있기 때문이다. 결과적으로, 이미지 세분화 작업에는 세분화에 특화된 모델을 사용하는 것이 더욱 적합하다는 점을 확인할 수 있었다.

이번 연구에서 제안된 시스템은 단순히 취약 계층의 밥 섭취량을 측정하는 것에 그치지 않고, 다양한 임상 데이터와 연계하여 보다 포괄적인 영양 모니터링 시스템으로 확장될 수 있다. 예를 들어, 사용자의 성별, 나이, 건강 상태와 같은 정보를 기반으로 맞춤형 영양 분석을 제공할 수 있다. 향후 연구에서는 식품군별 분석을 확대할 필요가 있다. 채소, 과일, 단백질 등 다양한 식품군에 따른 잔량 분할 및 영양소 분석 방법을 개발하면, 보다 정밀한 식사 모니터링과 맞춤형 영양 제공이 가능해질 것이다. 이를 통해 취약 계층의 건강 상태를 효과적으로 관리하고 개선할 기회를 제공할 수 있다. 따라서 이 시스템은 취약 계층의 영양 관리 문제를 해결하는 데 핵심적인 도구로 기여할 것으로 기대된다.

References

- [1] G. Hong, H. Jang, Y. Shin, "Predictors of malnutrition among community-residing older adults in Korea", *The Gerontologist*, vol. 55, Suppl. 2, November 2015, pp. 124-124, doi: 10.1093/geront/gnv508.02.
- [2] E. Dent, O. Hoogendijk, R. Visvanathan, O. R. L. Wright, "Malnutrition Screening and Assessment in Hospitalized Older People: A Review", *The Journal of Nutrition, Health & Aging*, vol. 23, no. 3, May 2019, pp. 431-441, doi: 10.1007/s12603-019-1176-z.
- [3] S. W. Ham, K. H. Kim, "Evaluation of the Dietary Quality and Nutritional Status of Elderly People Using

- the Nutrition Quotient for Elderly (NQ-E) in Seoul”, *Journal of Nutrition and Health*, vol. 53, no. 1, Januar y 2020, pp. 68-82, doi: 10.4163/jnh.2020.53.1.68.
- [4] J. Lubura, L. Pezo, M. A. Sandu, V. Voronova, F. Donsì, J. Šic Žlabur, “Food Recognition and Food Waste Estimation Using Convolutional Neural Network”, *Electronics*, vol. 11, no. 22, November 2022, pp. 1-16, doi: 10.3390/electronics11223746.
- [5] Y. A. Sari, E. N. Candra, S. Adinugroho, F. Utamingrum, J. M. Maligan, N. Nur'aini, “Recognition of Le flover Food Based on Deep Learning”, *Advances in Automation, Mechanical and Design Engineering*, December 16-18, 2022, Beijing, China, pp. 207-215, doi: 10.1007/978-3-031-40070-4_16.
- [6] W. Zhang, D. Zhao, W. Gong, Z. Li, Q. Lu, S. Yang, “Food Image Recognition with Convolutional Neural Networks”, 2015 IEEE UIC-ATC-ScalCom Conference, August 10-14, 2015, Taipei, Taiwan, pp. 520-525, doi: 10.1109/UIC-ATC-ScalCom-CBDCCom-IoP.2015.139.
- [7] M. M. Anthimopoulos, L. Gianola, L. Scarnato, P. Diem, S. G. Mougiakakou, “A Food Recognition System for Diabetic Patients Based on an Optimized Bag-of-Features Model”, *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, July 2014, pp. 1261-1271, doi: 10.1109/JBHI.2014.2308928.
- [8] Y. He, C. Xu, N. Khanna, C. J. Boushey, E. J. Delp, “Food Image Analysis: Segmentation, Identification and Weight Estimation”, 2013 IEEE International Conference on Multimedia and Expo (ICME), July 15-19, 2013, San Jose, CA, USA, pp. 1-6, doi: 10.1109/ICME.2013.6607548.
- [9] K. Roy, S. S. Chaudhuri, S. Pramanik, “Deep Learning Based Real-Time Industrial Framework for Rotten and Fresh Fruit Detection Using Semantic Segmentation”, *Microsystem Technologies*, vol. 27, no. 9, November 2020, pp. 3365-3375, doi: 10.1007/s00542-020-05123-x.
- [10] X. Wu, S. Yu, E. P. Lim, C. W. Ngo, “OVFoodSeg: Elevating Open-Vocabulary Food Image Segmentation via Image-Informed Textual Representation”, *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 17-21 2024, Seattle, USA, pp. 4144-4153, doi: 10.1109/CVPR52733.2024.00397.
- [11] Z. Pan, J. Xu, Y. Guo, Y. Hu, G. Wang, “Deep Learning Segmentation and Classification for Urban Village Using a Worldview Satellite Image Based on U-Net”, *Remote Sensing*, vol. 12, no. 10, May 2020, pp. 1-17, doi: 10.3390/rs12101574.
- [12] O. Ronneberger, P. Fischer, T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, October 5-9, 2015, Munich, Germany, pp. 234-241, doi: 10.1007/978-3-319-24574-4_28.
- [13] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, “Attention U-Net: Learning Where to Look for the Pancreas”, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, May 14-18, 2018, Istanbul, Turkey, pp. 29-37, doi: 10.48550/arXiv.1804.03999.
- [14] D. Maji, P. Sigedar, M. Singh, “Attention Res-UNet with Guided Decoder for semantic segmentation of brain tumors”, *Biomedical Signal Processing and Control*, vol. 71, January 2022, pp. 1-10, doi: 10.1016/j.bspc.2021.103077.
- [15] F. I. Diakogiannis, F. Waldner, P. Caccetta, C. Wu, “ResUNet-a: A Deep Learning Framework for Semantic Segmentation of Remotely Sensed Data”, *Remote Sensing*, vol. 12, no. 5, April 2020, pp. 94-114, doi: 10.1016/j.isprsjprs.2020.01.013.