

Enhanced White Blood Cell Classification via YOLO-based Region-of-Interest Extraction and Feature Ensemble

Junhwi Park*, Changjoon Park*, Namjung Kim**, Ryumduck Oh***, Jeonghwan Gwak***

*Student, Dept. of IT-Energy Convergence, Korea National University of Transportation, Chungju, Korea

**Student, Dept. of Software, Korea National University of Transportation, Chungju, Korea

***Professor, Dept. of Software, Korea National University of Transportation, Chungju, Korea

[Abstract]

Blood is composed of white blood cells (WBC), red blood cells (RBC), and platelets, components closely related to human health. They respectively contribute to the immune system, the transfer of oxygen to body organs, and bleeding prevention. In the case of WBC, when abnormalities occur in the body, such as infections and allergic reactions, the relevant type of WBC in the blood increases to respond. Therefore, it is possible to identify the type of WBC in the case of abnormalities in the body through the application of deep learning-based object detection algorithms and classification methods. This information can be used to predict the health status of patients easily. Therefore, this paper proposes a method for extracting the region of interest (RoI) for WBC and classifying WBC types based on blood microscope images using you-only-look-once (YOLO) and feature ensemble techniques. We select RoI extraction models through a comparative analysis of WBC detection performance for YOLO V5, V8, V9 and YOLO-Neural Architecture Search (YOLO-NAS), and demonstrate an improvement in WBC type classification performance through comparative analysis and a Top-3 model feature ensemble based on general convolutional neural network (CNN) models, such as ResNet and EfficientNet. Compared to the approximately 98% performance of the Top-3 models based on the F1-Score, it achieved an improved performance of approximately 99%.

▶ **Key words:** Heterogeneous Feature Ensemble, Homogeneous Feature Ensemble, Image Classification, Object Detection, You-Only-Look-Once (YOLO), White Blood Cell

-
- First Author: Junhwi Park, Corresponding Author: Jeonghwan Gwak
 - *Junhwi Park (objectdetection@kakao.com), Dept. of IT-Energy Convergence, Korea National University of Transportation
 - *Changjoon Park (cjp128@a.ut.ac.kr), Dept. of IT-Energy Convergence, Korea National University of Transportation
 - **Namjung Kim (knj95@kakao.com), Dept. of Software, Korea National University of Transportation
 - ***Ryumduck Oh (rdoh@ut.ac.kr), Dept. of Software, Korea National University of Transportation
 - ***Jeonghwan Gwak (jgwak@ut.ac.kr), Dept. of Software, Korea National University of Transportation
 - Received: 2024. 06. 13, Revised: 2024. 12. 31, Accepted: 2025. 01. 01.

[요 약]

혈액은 신체 건강과 밀접한 연관이 있으며, 백혈구(White Blood Cells, WBC), 적혈구(Red Blood Cells, RBC), 그리고 혈소판(Platelets)으로 구성된다. 이는 면역 체계, 신체 기관에 산소 전달, 지혈 작용 등에 기여한다. 백혈구의 경우 감염, 알레르기 반응과 같이 신체 내에 이상이 발생하면 해당 이상 증세와 관련된 백혈구 유형이 증가하는 특징을 가진다. 이러한 특징을 이용하면 딥러닝 기반 객체 탐지 알고리즘 및 분류 방법을 적용하여 백혈구 세부 유형을 식별할 수 있으며, 환자의 건강 상태 예측에 용이한 정보로 활용할 수 있다. 따라서, 본 논문은 You-Only-Look-Once (YOLO) 및 특징 앙상블 기반 혈액 현미경 이미지 내 관심 영역 추출 및 백혈구 세부 유형 분류 기법을 제안한다. YOLO V5, YOLO V8, YOLO V9, YOLO-Neural Architecture Search (YOLO-NAS) 모델의 백혈구 탐지 성능을 비교 분석하여 가장 적합한 관심 영역 추출 모델을 선정하였다. 이후 ResNet, EfficientNet과 같은 일반적인 합성곱 신경망(Convolutional Neural Network)의 백혈구 유형 분류 성능을 비교 분석한 후, 성능이 우수한 상위 3개의 모델을 선택하여 특징 앙상블을 진행하였으며 F1-Score를 기준으로 상위 3개 모델의 약 98%와 비교하여 약 99%의 향상된 성능을 달성하였다.

▶ **주제어:** 이중 특징 앙상블, 동종 특징 앙상블, 이미지 분류, 객체 탐지, 옰로, 백혈구

I. Introduction

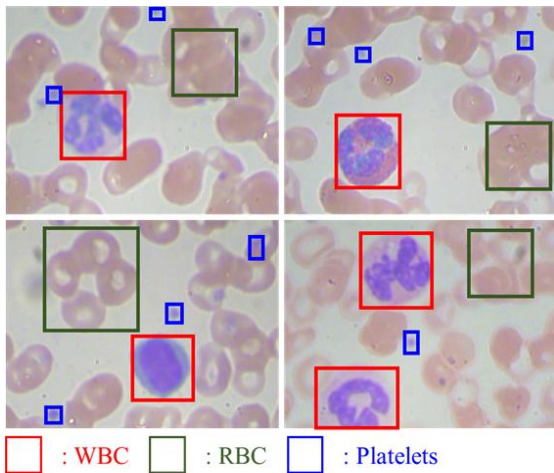


Fig. 1. Example of Blood Microscopic Images

혈액은 각 신체 기관에 산소를 공급하고 체온 조절 및 면역 체계 구성에 매우 중요한 역할을 한다. 혈액은 백혈구(White Blood Cells, WBC), 적혈구(Red Blood Cells, RBC), 혈소판(Platelets) 등의 다양한 세포로 구성되어 있으며 해당 요소 중 백혈구는 인체의 면역 체계의 주요 구성요소이다. 적혈구는 Figure 1의 초록색 박스(Box) 영역에서 확인할 수 있으며, 혈액 내 가장 큰 비중을 차지한다. 적혈구는 신체 각 기관에 산소를 전달하거나 체온 유지의 역할을 한다. 혈소판의 경우 Figure 1의 파란색 박스 영역에서 확인할 수 있으며, 지혈을 담당한다.

백혈구는 Figure 1의 빨간색 박스 영역에서 확인할 수 있으며 호중구(Neutrophil), 림프구(Lymphocyte), 단핵구(Monocyte), 호산구(Eosinophil)와 같은 유형으로 분류(Classification)된다. 호중구는 가장 흔한 유형의 백혈구로, 염증과 감염에 반응하여 병원체를 제거하는 역할을 한다. 림프구는 T 세포, B 세포, 자연 살해 세포(Natural Killer Cell, NK Cell)로 구성되어 있으며, 감염 및 면역, 종양 제어에 중요한 역할을 한다. 단핵구는 가장 큰 백혈구로, 조직으로 이동하여 특정 감염에 대항한다. 호산구는 주로 알레르기 반응이나 기생충 감염에 대한 면역 반응과 관련되며 염증 반응, 천식, 피부염 등과 연관된다.

신체 내에 바이러스 감염, 알레르기 또는 종양과 같이 면역 체계 저하로 인한 질병이 발생하면, 혈액 내에서 해당 질병과 관련된 유형의 백혈구가 증가한다. 따라서, 혈액 내에 증가한 백혈구 유형 분류 및 분석을 통해 환자의 건강 상태를 예측할 수 있으며, 이를 통해 특정 질병에 대한 사전 예방과 같은 지속적인 건강 관리에 중요한 지표로써 활용할 수 있다. Figure 1과 같이 혈액 현미경 이미지(Image)는 적혈구, 혈소판 등 다양한 혈액 구성요소를 포함하고 있다. 이와 같은 혈액 현미경 이미지의 경우, 백혈구 클래스(Class)가 차지하는 비율은 적혈구 및 혈소판보다 상대적으로 적은 비율을 차지한다. 강인한 성능을 가지는 백혈구 세부 유형 분류 모델(Model)을 설계하기 위해서는 각 유형에 대해 강인한 특징 추출이 필수적이다. 하

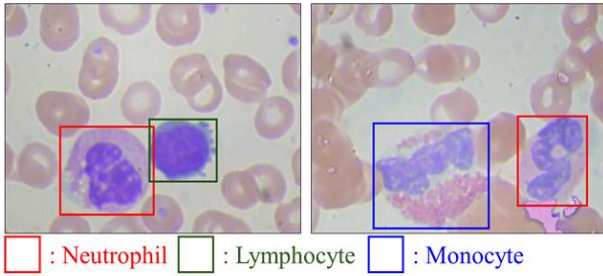


Fig. 2. Example of Data Showing Different Types of White Blood Cells in a Single Image

지만 적혈구, 혈소판으로 인해 형성되는 복잡한 배경은 백혈구 세부 유형 분류를 위한 특징 추출에 방해 요인이 될 수 있다. 또한, Figure 2와 같이 혈액 현미경 이미지의 경우 단일 이미지 내에 두 가지 이상의 유형에 해당하는 백혈구를 포함할 수 있다. 이와 같은 이미지는 일반적인 분류 알고리즘(Algorithm)을 통해 백혈구 세부 유형 분류가 어렵다는 한계를 가진다. 따라서, 본 논문에서는 복잡한 배경을 가지는 혈액 현미경 이미지 내에서 백혈구의 영역 특징을 극복해야 할 문제로 선정하였다.

위와 같은 한계점을 극복하기 위해 본 논문에서는 You-Only-Look-Once (YOLO)[1] 기반 객체 탐지 알고리즘을 활용하여 백혈구에 대한 관심 영역(Region of Interest, RoI)을 추출하고, 특징 앙상블(Feature Ensemble) 기반의 분류 작업을 통해 해당 문제를 극복할 수 있음을 입증한다.

먼저, 관심 영역 추출에 가장 적합한 YOLO 기반 객체 탐지 알고리즘을 선정하기 위해 Figure 1과 같은 양상을 가지는 객체 탐지 데이터 세트(Dataset)를 구성하였으며 해당 데이터 세트를 통해 추출한 관심 영역을 기반으로 분류 데이터 세트를 구성하였다. 각 데이터 세트 목적에 따라 해당 데이터 세트에서 가장 강인한 성능을 가지는 모델을 선정하기 위해 YOLO 기반 객체 탐지 알고리즘의 객체 탐지 성능 실험 및 일반적으로 분류 성능이 강인하다고 알려진 ResNet[2], DenseNet[3] 등과 같은 일반적인 합성곱 신경망(Convolutional Neural Network)의 백혈구 세부 유형 분류 실험을 진행하였다. 두 번째로, 각 실험 결과에 대해 정량적·정성적 성능평가를 진행하여 백혈구 관심 영역 추출에 가장 적합한 YOLO 기반 객체 탐지 알고리즘을 선정하고 백혈구 세부 유형 분류에서 강인한 성능을 가지는 Top-3 모델을 선정한다. 마지막으로, 선정된 Top-3 모델의 특징 앙상블을 통한 백혈구 세부 유형 분류 기법을 제안한다. 본 논문의 기여도는 다음과 같다.

- YOLO 기반 객체 탐지 알고리즘의 혈액 현미경 이미지 내 백혈구에 대한 강인한 객체 탐지 성능을 도출함으로써, YOLO 기반 객체 탐지 알고리즘이 혈액 구성요소 객체 탐지에서 유의미한 성능을 가짐을 확인한다.
- Top-3 합성곱 신경망 모델로 ConvNeXt[4] Large, ConvNeXt Small, ResNet-152를 선정함으로써 백혈구 세부 유형 분류에 강인한 특징 추출에 특화된 합성곱 신경망을 확인하였으며 정량적, 정성적 평가를 통한 체계적인 모델 설계 방안을 제시한다.
- 선정한 Top-3 모델의 특징 앙상블을 통해 최종적으로 일반적인 합성곱 신경망 모델보다 강인한 성능을 가지는 백혈구 세부 유형 분류 모델을 제안하며 두 번째 기여에서 언급한 모델 설계 방안의 근거를 정성적 평가를 통해 제시한다.

최종적으로, YOLO 기반 객체 탐지 알고리즘이 혈액 현미경 이미지에서 유의미한 성능을 가짐을 확인하였으며, 객체 탐지 알고리즘과 특징 앙상블 기법을 활용하여 기존 일반적인 합성곱 신경망보다 강인한 성능을 보이는 백혈구 세부 유형 분류 모델을 제안한다.

본 논문의 구성은 I. Introduction, II. Preliminaries, III. The Proposed Method, IV. Experiments, V. Conclusions and Future Works 총 다섯 개의 장으로 구성된다. I 장에서는 백혈구 세부 유형 분류의 필요성 및 본 논문의 기여도에 대해 언급한다. II 장에서는 기존에 진행된 백혈구 관련 연구에 대해 분석하고 기존 연구와 본 논문에서 제안하는 기법이 가지는 차이점을 언급한다. 또한, 본 연구의 이해를 돕기 위한 배경지식을 소개한다. III 장에서는 본 논문에서 제안하는 기법 상세와 이해를 돕기 위한 배경지식을 언급한다. IV 장에서는 체계적인 실험을 통해 관심 영역 추출을 위한 객체 탐지 모델 및 특징 앙상블을 위한 합성곱 신경망 모델을 선정한 근거를 제시한다. 마지막 V 장에서는 최종적인 결론과 본 저자의 향후 연구 계획을 언급하며 본 논문을 마무리한다.

II. Preliminaries

1. Image Classification

백혈구와 같은 혈액 현미경 이미지에서의 세포 분류 연구는 과거부터 현재까지 활발하게 진행되고 있다. M. C.

Su, et al.[5]의 연구에서는 혈액 현미경 이미지에 대해 지역 방향 패턴(Local Directional Pattern, LDP) 및 색상 특징 등으로 구성된 20가지의 특징을 추출한다. 추출한 특징에 대해 신경망을 활용하여 백혈구 세부 유형 분류에 있어 99.11%의 높은 성능을 보였다. 하지만, 백혈구 영역에 대해서만 특징을 추출하는 것이 아니기에 배경과 같은 다른 요인으로 인해 잡음을 포함할 가능성이 있다. 따라서, 본 논문에서는 백혈구의 온전한 영역 추출을 위해 바운딩 박스(Bounding Box) 기반 관심 영역 추출을 통해 백혈구 영역 손실에 대한 한계를 극복하였다.

O. Saidani, et al.[6]의 연구에서는 백혈구 세부 유형 분류를 위해서 혈액 현미경 이미지 전처리 방법론 및 새로운 구조를 가지는 합성곱 신경망을 제안했다. 혈액 현미경 이미지 전처리 방법은 이미지 크기 재조정, Red, Green, Blue (RGB)의 3채널(Channel)을 가지는 이미지에 대해 luma component, red projection, and blue projection (YUV) 변환 적용, Histogram Equalization 및 Edge Smoothing 기법 적용, RGB 채널로 이미지 재변환의 단계로 구성된다. 해당 과정을 통해 혈액 현미경 이미지의 잡음을 제거할 수 있으며 제안하는 합성곱 신경망의 계산 비용 또한 절감할 수 있다. 새로운 구조를 가지는 합성곱 신경망의 경우 3개의 합성곱 신경망과 3개의 다층 신경망으로 구성되었으며, 구성요소 사이에 Dropout을 적용한 형태를 가진다. 이와 같은 이미지 전처리 및 제안한 합성곱 신경망을 통해 강인한 백혈구 세부 유형 분류 정확도를 달성하였다. 하지만, 혈액 현미경 이미지 내에 다른 유형을 가지는 백혈구가 2개 이상 존재하는 경우 각 백혈구에 대한 세부 유형 분류는 어렵다는 한계를 가진다.

2. Object Detection

세포와 같이 작은 객체에 관한 탐지 연구 현황은 다음과 같다. T. Wollmann, et al.[7]은 현미경 이미지 내 세포 및 입자 감지를 위한 Deep Consensus Network를 제안하였으며 Feature Pyramid Network, Centroid Proposal Network, Anchor를 활용하여 세포 탐지를 수행하였다. 또한, 정규화 Anchor 및 상호 정보 기반의 새로운 손실함수 도입, Non-Maximum Suppression (NMS) 알고리즘 개선을 통해 좋은 결과를 달성하였다.

M. L. Ali, et al.[8]의 연구에서는 림프구 분석을 위한 채널 강화 전이 학습 기반 Channel Boosted Hybrid Vision Transformer (CB-HVT)를 제안했다. CB-HVT는 Transformer와 합성곱 신경망을 결합하였으며, 다양한 모듈(Module)로 구성되어, 강화된 채널 생성, 병합, 영역

인식 등을 통해 효과적인 림프구 인식률을 달성하였다. CB-HVT는 타 모델에 비해 강건한 성능을 보이는 것을 입증하였으며, 해당 모델을 임상 병리학 분야에 상용화 가능성을 확인하였다.

X. Zheng, et al.[9]에서는 백혈구 검출을 위해 두 가지 단계로 구성된 객체 탐지 모델을 제안했다. 첫 번째 단계에서는 입력된 혈액 현미경 이미지 내에서 Adaptive Center-Surround Difference (ACSD) 알고리즘을 기반으로 백혈구에 관한 관심 영역을 추출한다. 두 번째 단계에서는 해당 관심 영역에 대해 CenterNet을 통한 백혈구 객체 탐지를 진행한다. YOLO 기반 객체 탐지 알고리즘의 경우 이미지 내 관심 객체에 관한 관심 영역 제안을 위한 기법으로는 일반적으로 Anchor Box가 사용된다. 하지만 해당 연구에서는 관심 영역 제안에 Anchor Box를 사용하지 않는 CenterNet을 사용하였다. 해당 기법을 통해 컴퓨터(Computer) 연산량 측면에서 Anchor 기반 관심 영역 제안 모델보다 효율적이고 더 정확한 백혈구 탐지 성능을 확인하였다.

3. Segmentation

혈액 현미경 이미지 내 백혈구 영역 분할 연구 또한 진행되었다. A. Abdurrazzaq, et al.[10]의 연구에서는 혈액 세포 이미지 내 백혈구 분할을 위해 기존에 사용된 U-Net[11]과 같은 일반적인 이미지 분할 기법과 달리 색상 정보를 바탕으로 수학적 접근 방식을 통해 RGB 이미지 내 백혈구 분할 방법을 제안했다. 해당 방법을 통해 이미지 분할 알고리즘에 정확도 및 효율성 개선에 효과가 있음을 확인했다.

혈액 현미경 이미지 내의 백혈구의 경우 세포질(cytoplasm)과 핵(nuclei)으로 구성된다. Y. Guo, et al.[12]의 연구에서는 3종류의 혈액 현미경 이미지 내에서 백혈구의 영역을 세포질과 핵으로 분할하는 기법을 제안한다. 혈액 현미경 이미지에서 백혈구 분할 정확도 향상을 위해 Indeterminacy fusion mechanism을 적용한 인코더(Encoder)-디코더(Decoder)의 구조를 가지는 합성곱 신경망을 제안했다. Indeterminacy fusion mechanism이란 복잡한 배경을 가지는 혈액 현미경 이미지 내에서 백혈구 영역을 정확하게 분할 하기 위해서 각 픽셀(Pixel)이 가지는 불확실성 특징을 활용하는 기법이다. 해당 기법을 통해 관심 객체(백혈구)와 배경을 더 정확하게 구분할 수 있으며 기존에 이미지 분할을 위해 제안된 U-Net과 같은 합성곱 신경망 분할 알고리즘보다 강인한 백혈구 분할 성능에 도달하였다.

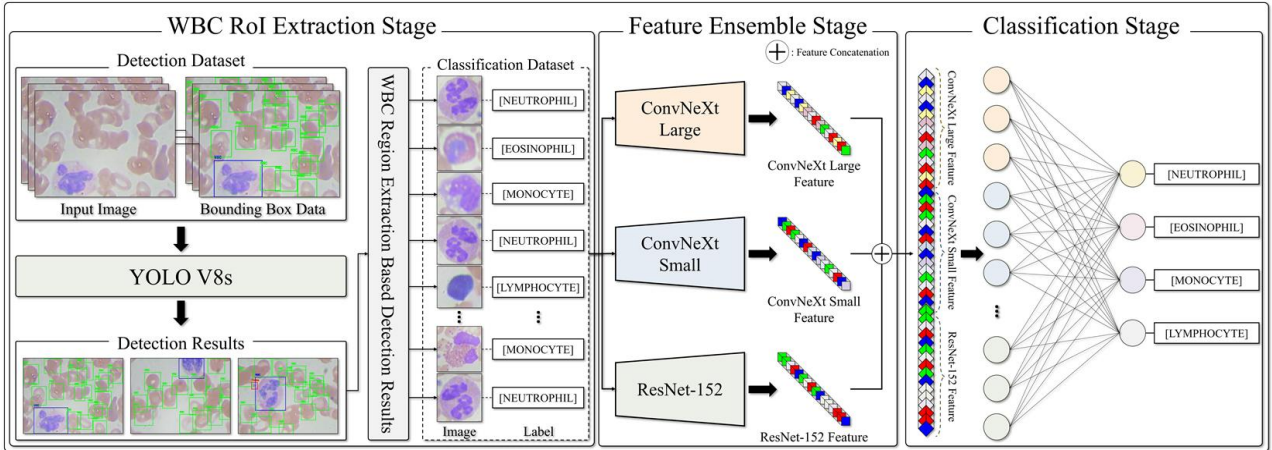


Fig. 3. Proposed Model Architecture

백혈구는 혈액 현미경 이미지를 통해 세부 유형 분류, 객체 탐지, 객체 분할 등 다양한 인공지능 분야에서 연구된다. 본 논문에서는, 상기 언급한 기존 연구의 한계를 극복하기 위해 혈액 현미경 이미지 내에서 백혈구의 관심 영역을 추출하고 특징 앙상블을 활용하여 백혈구 세부 유형 분류에 강인한 모델을 제안한다.

4. Background Knowledge

합성곱 신경망은 이미지 내 특징 추출을 위해 필터(Filter), 풀링(Pooling), 정규화(Normalization), 활성화 함수(Activation Function) 등의 다양한 요소로 구성된다. 필터는 필터의 크기, 필터의 적용 간격(Stride), 필터의 팽창 정도(Dilation)에 따라 다양한 종류의 필터가 존재한다. 풀링은 최대 풀링(Max Pooling), 평균 풀링(Average Pooling) 등 다양한 종류의 기법이 있으며, 정규화 기법은 적용 범위에 따라서 배치 정규화(Batch Normalization), 레이어 정규화(Layer Normalization), 객체 정규화(Instance Normalization), 그룹 정규화(Group Normalization)[13]처럼 다양한 기법이 존재한다. 활성화 함수 또한 Rectified Linear Unit (ReLU), Sigmoid Linear Unit (SiLU)과 같이 다양한 종류가 있다. 합성곱 신경망의 구성요소가 다양한 만큼 합성곱 신경망 모델들은 구성요소의 조합에 따라 각기 다른 양상의 특징 맵(Feature Map)을 추출한다. 특징 앙상블 기법은 여러 합성곱 신경망이 추출한 특징 맵을 다층 신경망의 입력으로 이용함으로써 결과를 판단하는 데에 다양한 특징을 활용할 수 있다는 장점을 가진다.

특징 앙상블 방법론은 크게 동종 특징 앙상블(Homogeneous Ensemble)과 이종 특징 앙상블(Heterogeneous Ensemble)로 구분된다. 동종 특징 앙상

블의 경우 동일한 모델 구조를 가지지만 깊이가 다른 합성곱 신경망 모델이 출력한 특징 맵을 이용하는 기법이다. 이와 달리, 이종 특징 앙상블의 경우 다른 구조를 가진 합성곱 신경망이 추출한 특징 맵을 이용하는 기법이다. 동종 특징 앙상블은 동일한 이미지에 대해 다양한 깊이의 특징을 이용할 수 있다는 장점이 있으며, 이종 특징 앙상블은 동일한 이미지에 대해 각기 다른 모델이 추출한 특징을 이용하기에 다양한 특징을 이용할 수 있다는 장점을 가진다.

III. The Proposed Method

1. Proposed Model Architecture

Figure 3은 혈액 현미경 이미지 내 백혈구 관심 영역 검출 및 유형 세부 분류를 위한 제안 모델 전체 구성을 도식화한 그림이다. 본 논문에서 제안하는 기법은 1) 백혈구 관심 영역 추출 단계(WBC RoI Extraction Stage), 2) 특징 앙상블(Feature Ensemble Stage) 단계, 3) 분류(Classification Stage) 단계 세 가지로 구성된다.

1) 백혈구 관심 영역 추출 단계에서는 백혈구 관심 영역 추출을 위해 Common Objects in COntext (COCO)[14]와 같은 대용량 데이터 세트에서 강인한 성능을 가지는 YOLO 기반 객체 탐지 알고리즘의 백혈구 탐지 성능에 대해 비교·분석하며 백혈구 관심 영역 추출에 가장 적합한 객체 탐지 알고리즘을 선정한다.

2) 특징 앙상블 단계에서는 Top-3 합성곱 신경망 모델이 분류기의 입력으로 사용하는 특징 맵을 추출하는 과정을 의미한다. Top-3 합성곱 신경망 모델로는 ConvNeXt Large, ConvNeXt Small, ResNet-152를 선정하였다. 모델 선정을 위해서 16개의 일반적인 합성곱 신경망 모델의

Table 1. BCCD Dataset (Used in the Experiment)

Detection Dataset		
Images	Image Size	Class Type
660	640×480	RBC, WBC, Platelets
Classification Dataset		
Images	Image Size	Class Type
8,000	64×64	Eosinophil, Lymphocyte, Monocyte, Neutrophil

백혈구 세부 유형 분류 실험을 진행하였다. Top-3 합성곱 신경망 모델은 각각 백혈구 세부 유형 분류에 가장 강인한 성능을 가지는 모델로 선정하였기에 해당 모델이 추출한 특징 맵은 백혈구 세부 유형 분류에서 유리한 특징임을 의미한다. 해당 특징들을 모두 활용하여 백혈구 세부 유형 분류를 진행하는 특징 앙상블을 통해 강인한 성능을 가지는 백혈구 세부 유형 분류 모델을 설계한다.

3) 분류 단계에서는 선정된 Top-3 합성곱 신경망 모델이 추출한 3가지 특징 맵의 입력을 통해 다양한 특징을 기반으로 백혈구 세부 유형을 분류하는 분류기(Classifier)를 설계하고, 기존 일반적인 합성곱 신경망 모델보다 강인한 백혈구 유형 분류 성능을 달성하였다.

따라서, 본 논문에서는 백혈구 세부 유형 분류 성능이 가장 높은 Top-3 합성곱 신경망 모델을 선정하고 3개 모델이 추출한 각기 다른 특징을 분류기의 입력으로 사용함으로써, 기존 성능 대비 더 강인한 백혈구 세부 유형 분류 성능 향상을 목표로 한다.

IV. Experiments

1. Dataset

혈액 현미경 이미지 데이터 세트는 Blood Cell Count and Detection (BCCD)[15]를 활용하였다. 본 데이터 세트는 혈액 현미경 이미지 내 혈액 구성요소 객체 탐지를 위한 데이터 세트와 백혈구 세부 유형 분류를 위한 데이터 세트로 구성된다. 객체 탐지를 위한 데이터 세트는 기존 410장이 제공되었지만, 바운딩 박스 정보가 누락된 67장을 제외한 343장을 활용하였다. 이미지 분류를 위한 데이터 세트는 호중구, 림프구, 단핵구, 호산구 총 4가지 클래스를 가지며 객체 탐지를 위한 데이터 세트의 이미지 높이와 너비를 절반으로 잘라 1장의 이미지를 4장으로 재구성하였다. 해당 데이터에 대해 회전, 반전 등 다양한 데이터 증강 기법을 적용하여 최종적으로 4가지 백혈구 유형마다 약 3,100장으로 구성한 것을 확인하였다. 본 논문에서는 혈액 현미경 이미지 내 강인한 백혈구 세부 유형 분류 모

델 제안을 위해 1) 백혈구 객체 탐지 실험과 2) 백혈구 유형 분류 실험을 설계하였다. 실험 진행을 위한 데이터 세트 구성은 Table 1과 같다.

먼저, 백혈구 관심 영역 추출을 위한 1) 백혈구 객체 탐지 실험 데이터 세트는 Table 1의 Detection Dataset 항목에서 확인할 수 있다. 기존 백혈구 유형 분류를 위한 BCCD 데이터 세트는 원본 이미지 데이터에 데이터 증강 기법을 적용한 형태로 구성된다. 따라서, 증강된 데이터 세트에도 강인한 객체 탐지 성능을 위해서 기존 BCCD 데이터 세트에 회전 및 반전과 같은 데이터 증강 기법을 적용하여 총 660장으로 객체 탐지 데이터 세트를 구성하였다. 학습 데이터의 경우 500장, 검증 데이터의 경우 100장, 테스트 데이터의 경우 증강을 적용하지 않은 데이터 60장으로 설정하였다. 다음으로, 이미지 분류를 위한 데이터 세트는 Table 1의 Classification Dataset 항목에서 확인할 수 있다. 이는 기존 BCCD 데이터 세트에 포함된 이미지 분류를 위한 데이터 세트에 객체 탐지 알고리즘을 적용하여 추출한 관심 영역을 통해 총 8,000장으로 구성하였다. 실험을 위한 데이터 세트는 학습 데이터 4,800장, 검증 데이터 2,000장, 테스트 데이터 1,200장으로 구성하였다. 데이터 세트 구성 비율은 클래스마다 2,000장으로 동일한 비율을 가지도록 하였다. 백혈구 세부 유형 분류를 위한 데이터 세트인 백혈구 관심 영역의 경우 Figure 1에서 확인할 수 있는 것처럼 640 × 480 크기를 가지는 이미지 내에 일부 영역을 차지한다. 따라서, 과도한 이미지 크기 조정은 이미지 내 객체의 특징을 파괴할 가능성을 가지기에 이러한 현상을 방지하고자, 백혈구 세부 유형 분류 실험의 경우 추출된 관심 영역의 크기와 큰 차이가 없도록 64 × 64로 재조정하여 진행하였다.

2. Experiment Results

2.1 Evaluation Metrics

본 논문에서는 혈액 현미경 이미지 내 백혈구 관심 영역 추출을 위한 객체 탐지 알고리즘 및 백혈구 세부 유형 분류 성능 비교 분석에 대해서 정량적·정성적 성능평가를 진행한다. 정량적 평가를 위한 평가지표는 정밀도

(Precision), 재현율(Recall), 정확도(Accuracy), F1-Score, mean Average Precision (mAP)를 선정하였으며, 정성적 평가를 위한 평가지표는 각 객체 탐지 알고리즘의 테스트 데이터 세트에 대한 백혈구 유형 바운딩 박스 예측 결과에 대한 성능으로 선정하였다. 정밀도, 재현율, 정확도에 대한 수식은 다음 수식(1)-(3)과 같다.

$$Precision = \frac{TP}{TP + FP}, \quad (1)$$

$$Recall = \frac{TP}{TP + FN}, \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}. \quad (3)$$

수식(1)-(3)의 T 는 True, F 는 False, P 는 Positive, N 은 Negative를 의미한다. T 와 F 는 모델이 예측한 결과에 대한 정답 여부를 의미하며 P 와 N 은 모델이 예측한 결과를 의미한다. 예를 들어, TN 은 모델이 Negative라고 예측한 결과가 정답임을 의미한다. 수식(1) Precision은 객체 탐지 모델이 양성이라고 예측한 샘플 중에 실제로 양성인 비율을 의미하며, 수식(2) Recall은 모델이 Ground Truth (GT)를 얼마나 잘 예측했는지에 대한 평가지표이다. 또한, 수식(3) Accuracy는 모델의 예측값과 실제값이 얼마나 일치하는지에 대한 지표이다. F1-Score의 수식은 다음과 같다.

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \quad (4)$$

수식(4) F1-Score는 Precision과 Recall의 Trade-off 관계에 따라 조화평균을 통해 두 가지 지표를 모두 고려할 수 있는 평가지표이다. 높은 F1-Score는 모델의 정밀도와 재현율 측면에서 모두 강인하다는 것을 의미한다. mAP는 클래스 당 AP의 평균을 의미하며, AP는 다음과 같다.

$$Average Precision(AP) = \int_0^1 P(R) dR. \quad (5)$$

수식(5)에서 P 는 정밀도를 의미하고, R 은 재현율을 의미한다. AP는 Precision과 Recall이 가지는 Trade-Off 관계를 시각화하여 Precision-Recall 곡면의 면적을 나타내기에 지표가 높을수록 객체 탐지 모델의 강인한 관심 영역 탐지 성능을 의미하는 평가지표이다. mAP는 클래스별 AP의 평균을 통해 산출할 수 있으며 수식은 다음과 같다.

$$mAP = \frac{1}{n} \sum_{c=1}^n AP_c. \quad (6)$$

수식(6)에서 n 은 클래스의 개수를 의미하고, c 는 클래스의 종류를 의미한다. 평가지표 수식(1)-(4), (6)을 통해 다양한 측면에서 성능 비교 분석을 진행하여, 연구 목적에

부합하는 최적의 알고리즘을 선정하였다.

또한, 객체 탐지 모델을 평가하기 위해 수식(6) mAP 지표는 모델이 예측한 바운딩 박스와 GT의 Intersection over Union (IoU)를 기준으로 도출한다. IoU는 다음과 같은 수식을 따른다.

$$IoU = \frac{Pred_{Bbox} \cap GT_{Bbox}}{Pred_{Bbox} \cup GT_{Bbox}}. \quad (7)$$

수식(7)의 $Pred_{Bbox}$ 는 모델이 예측한 바운딩 박스를 의미하며, GT_{Bbox} 는 바운딩 박스 정답을 의미한다. IoU는 0 이상 1 이하의 값을 가지는 평가지표로서 모델이 예측한 바운딩 박스 영역과 GT 바운딩 박스 영역이 정확하게 일치하는 정도를 의미한다. 객체 탐지 모델을 평가하기 위한 mAP 지표는 대표적으로 계산에 반영된 IoU에 따라 mAP_0.5와 mAP_0.5:0.95로 나뉜다. mAP_0.5의 경우 모델이 예측한 바운딩 박스와 GT의 IoU가 0.5 이상인 샘플에 대한 AP의 평균이며, mAP_0.5:0.95의 경우 IoU가 0.5 이상 예측 결과부터 점진적으로 0.95인 예측 결과의 평균을 통해 계산된다. 따라서, mAP_0.5의 경우 객체 탐지 모델의 보편적인 관심 영역 탐지 성능을 확인할 수 있으며, mAP_0.5:0.95를 통해 객체 탐지 모델이 예측한 바운딩 박스의 정밀한 정도를 확인할 수 있다.

2.2 WBC RoI Extraction Stage

혈액 현미경 이미지 내 백혈구 관심 영역 추출을 위한 YOLO 모델 선정을 위해 백혈구에 대한 탐지 성능의 비교·분석을 진행하였다. COCO 사전 학습 가중치를 사용하였으며, 과적합을 방지하기 위해 YOLO Loss를 기준으로 갱신이 50 Epoch 이상 진행되지 않는다면 학습을 종료하는 Early Stopping을 적용하였다. 본 논문에서 사용된 YOLO 모델은 혈액 현미경 이미지 내 백혈구 관심 영역 추출을 위한 모델이기에 백혈구에 대한 객체 탐지 성능이 우선시된다. 이에 mAP, Recall과 같이 객체 탐지 알고리즘을 위한 평가지표에 적혈구, 혈소판 같은 다른 혈액 구성요소에 대한 객체 탐지 성능이 반영된다면, 백혈구에 대한 객체 탐지 성능을 확인하는 것에 잡음으로 작용할 수 있다. 따라서, 본 논문에서 백혈구 관심 영역 추출을 위한 객체 탐지 알고리즘 평가지표에는 다른 혈액 구성요소에 대한 탐지 성능은 고려하지 않았다. 백혈구에 대한 YOLO 모델의 객체 탐지 성능은 Table 2와 같다. YOLO 모델이 추출한 백혈구 관심 영역은 Figure 3의 특징 앙상블 단계의 입력으로 활용되기에 정확하고 정밀한 객체 탐지 성능이 필수적으로 요구된다. 따라서, 모델 선정을 위해 정량

Table 2. Detection Models Performance Results

Class	Models	Precision	Recall	mAP_0.5	mAP_0.5:0.95
WBC	YOLO V8m	0.971	1.000	0.992	0.803
	YOLO V8s	0.928	1.000	0.995	0.804
	YOLO V9e	0.969	1.000	0.995	0.800
	YOLO V9c	0.969	0.984	0.991	0.811
	YOLO V5m	0.973	1.000	0.995	0.802
	YOLO V5s	0.973	1.000	0.995	0.787
	YOLO-NAS_M	0.917	0.985	0.998	0.794
	YOLO-NAS_L	0.905	1.000	0.999	0.800

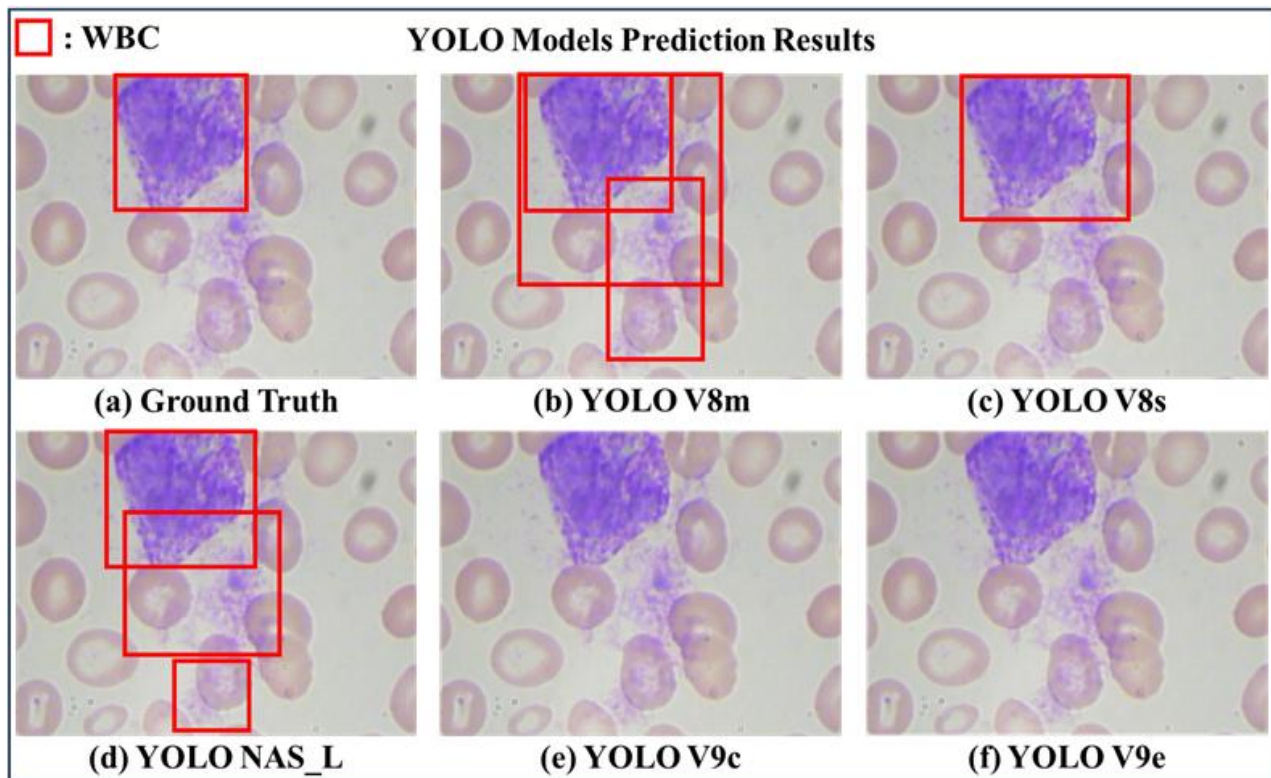


Fig. 4. YOLO Models Prediction Samples

적 성능평가의 기준이 되는 평가지표를 Recall과 mAP_0.5:0.95로 선정하였으며, 정성적 성능평가 기준으로는 테스트 데이터 세트에 대한 예측 결과 중 False Positive Sample (FP 샘플)의 발생 분석을 토대로 선정하였다. 본 논문에서는 백혈구 관심 영역 추출을 위한 객체 탐지 알고리즘으로 YOLO V8s를 선정하였다. 일반적으로 객체 탐지 알고리즘에서 예측한 수많은 바운딩 박스에서 가장 정답과 가까운 바운딩 박스를 선정하는 방법은 NMS 알고리즘을 이용한다. NMS 알고리즘은 모델이 예측한 바운딩 박스에 대해 클래스 예측 확률과 IoU Score를 임계치로 설정하여 설정한 임계치보다 낮은 값을 가지는 바운딩 박스는 소거하는 방식으로 작동한다. 모든 객체 탐지 알고리즘 비교군에 대해 바운딩 박스 예측 결과 도출을 위한 NMS 알고리즘의 하이퍼파라미터는 클래스 분류 정

확도 0.25, IoU Score 0.6 이상인 탐지 결과만 후처리하도록 설정하였다. YOLO V8s는 Table 2에서 볼 수 있듯이, Recall과 mAP_0.5:0.95 측면에서 가장 강인한 성능을 보이는 것을 확인할 수 있다. Figure 4는 동일한 테스트 데이터 세트 이미지에 대해 각 YOLO 모델이 예측한 바운딩 박스를 시각화한 결과이다. Figure 4의 (b)와 (d)는 각각 YOLO V8m과 YOLO NAS_L의 바운딩 박스 예측 결과를 시각화한 그림이다. 해당 그림에서 확인할 수 있는 것처럼, 두 모델은 일부 데이터 세트에 대해 백혈구에 대한 FP 샘플 바운딩 박스가 발생한 것을 확인할 수 있다. 일반적으로, 합성곱 신경망 기반 알고리즘은 깊이가 더 깊고, 넓을수록 강인한 성능을 가진다고 알려져 있다. 하지만, Figure 4의 (c)에서 확인할 수 있는 YOLO V8s의 예측 결과를 보면 신경망 깊이가 상대적으로 더 깊은

Table 3. Classification Performance Results (Weight Update Disabled)

Models	Precision	Recall	Accuracy	F1-Score
AlexNet	0.9033	0.9033	0.9517	0.9033
ConvNeXt Large	0.8583	0.8583	0.9298	0.8583
ConvNeXt Small	0.7925	0.7925	0.8974	0.7925
DenseNet121	0.7950	0.7950	0.8989	0.7950
DenseNet201	0.8317	0.8317	0.9166	0.8317
EfficientNet_b0	0.7783	0.7783	0.8908	0.7783
EfficientNet_b1	0.7767	0.7767	0.8900	0.7767
EfficientNet_b4	0.6617	0.6617	0.8334	0.6617
EfficientNet_b7	0.6192	0.6192	0.8143	0.6192
RegNet	0.8567	0.8567	0.9289	0.8567
ResNet-50	0.7875	0.7875	0.8950	0.7875
ResNet-101	0.7942	0.7942	0.8986	0.7942
ResNet-152	0.8375	0.8375	0.9197	0.8375
VGG_11	0.9350	0.9350	0.9675	0.9350
VGG_16	0.9075	0.9075	0.9540	0.9075
VGG_19	0.8750	0.8750	0.9376	0.8750

Table 4. Classification Performance Results (Weight Update Enabled)

Models	Precision	Recall	Accuracy	F1-Score
AlexNet	0.9542	0.9542	0.9771	0.9542
ConvNeXt Large	0.9892	0.9892	0.9946	0.9892
ConvNeXt Small	0.9875	0.9875	0.9938	0.9875
DenseNet121	0.9608	0.9608	0.9804	0.9608
DenseNet201	0.9767	0.9767	0.9883	0.9767
EfficientNet_b0	0.9742	0.9742	0.9871	0.9742
EfficientNet_b1	0.9808	0.9808	0.9904	0.9808
EfficientNet_b4	0.9450	0.9450	0.9725	0.9450
EfficientNet_b7	0.9692	0.9692	0.9846	0.9692
RegNet	0.9792	0.9792	0.9896	0.9792
ResNet-50	0.9783	0.9783	0.9892	0.9783
ResNet-101	0.9817	0.9817	0.9908	0.9817
ResNet-152	0.9867	0.9867	0.9933	0.9867
VGG_11	0.9667	0.9667	0.9833	0.9667
VGG_16	0.9750	0.9750	0.9875	0.9750
VGG_19	0.9725	0.9725	0.9863	0.9725

YOLO V8m과 YOLO-NAS보다 강인한 백혈구 객체 탐지 성능을 확인할 수 있다. 또한, Figure 4의 (e)와 (f)에서 확인할 수 있는 동일한 데이터 세트에 대한 YOLO V9 기반 모델의 백혈구 탐지 예측 결과의 경우 백혈구를 탐지하지 못한 예측 결과를 확인할 수 있다. 백혈구의 경우 Figure 1과 Figure 4에서 확인할 수 있는 것처럼 적혈구와 혈소판보다 상대적으로 크고 보라색을 띠는 단순한 특징을 가진다. 복잡한 양상을 가지는 데이터 세트의 경우 이미지 내에 다양한 크기의 객체에 대해 다양한 특징 파악이 필요하기에 깊은 합성곱 신경망이 이미지의 특징을 추출하기 용이하다. 하지만, 단순한 양상을 가지는 데이터 세트에서의 과도하게 깊은 합성곱 신경망의 사용은 기울기 소실과 같은 문제점을 초래할 수 있으며 이에 따라 이미지의 강인한 특징 추출 또한 어렵다는 문제점을 가진다. 또한, FP 샘플은 Precision 지표의 향상 원인이기에 백혈구가 존재

하지 않는 영역에 대해서도 백혈구 영역이라고 예측할 가능성을 내포한다. 또한, Figure 4의 YOLO V9의 객체 탐지 결과인 (e)와 (f)의 경우 백혈구에 대해 아무것도 탐지하지 못하는 결과를 확인할 수 있다. 이와 같은 결과는 백혈구 세부 유형 분류 정확도를 저해하는 주요한 원인이 되기에 백혈구 관심 영역 추출을 위한 객체 탐지 알고리즘에서 제외하였다. 이러한 정량적·정성적 성능평가를 통해 YOLO V8s가 가장 백혈구 관심 영역 추출에 적합한 모델임을 확인하였으며, Figure 3의 백혈구 관심 영역 추출 단계를 위한 객체 탐지 알고리즘으로 선정하였다.

2.3 Feature Ensemble Stage

백혈구 세부 유형 분류를 위한 특징 앙상블 모델을 선정하기 위해 딥러닝 라이브러리 PyTorch를 활용하여 일반적인 합성곱 신경망 모델의 백혈구 세부 유형 분류 실험을

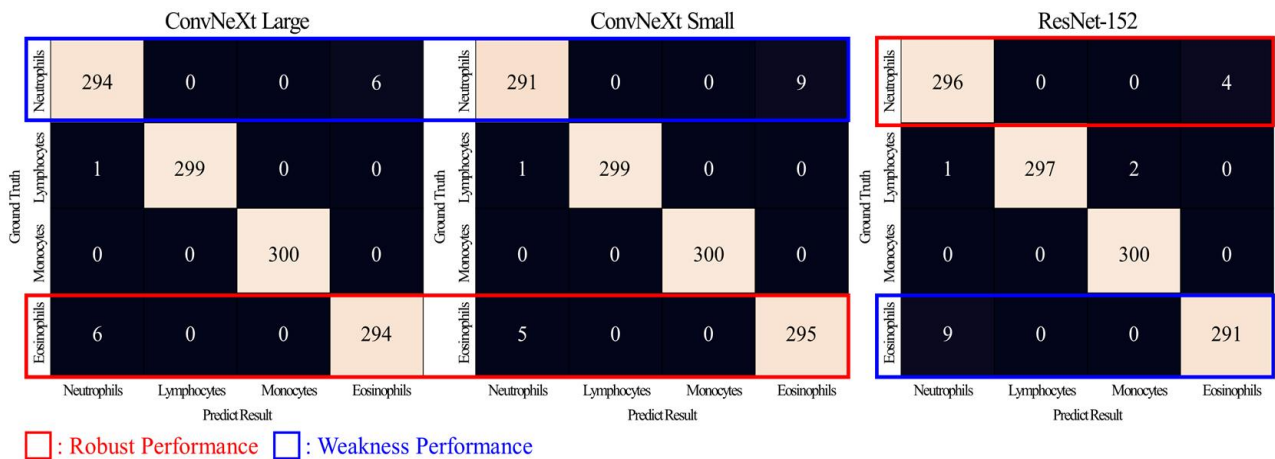


Fig. 5. Confusion Matrix Results(ConvNeXt Large, ConvNeXt Small, ResNet-152)

진행하였다. ImageNet[16]을 사전 학습한 가중치를 사용하였으며, 본 연구에서 객체 탐지 모델의 결과를 통해 구축한 분류 데이터 세트를 활용하여 모델의 가중치를 미세 조정하는 전이 학습을 진행하였다. 입력 이미지의 경우 64×64 크기를 가지도록 재조정하였으며, 각 픽셀의 RGB 채널에 대해 평균 : 0.485, 0.456, 0.406, 표준편차 : 0.229, 0.224, 0.225의 값을 가지도록 정규화를 진행하였다. 동일 모델에 대한 특징추출기의 가중치 갱신 활성화 여부에 따른 분류 성능을 비교하였다. 또한, 모델 과적합 방지 및 호중구, 림프구, 단핵구, 호산구로 구성된 4가지 유형의 백혈구를 분류하기에 F1-Score를 기준으로 7회 이상 성능이 향상되지 않으면 학습이 종료되는 Early Stopping을 적용하였다. 또한, 성능 지표 및 혼동행렬 분석을 통해 Figure 3의 2) 특징 앙상블 단계를 위한 Top-3 합성곱 신경망 모델을 선정하였다. Tables 3, 4는 각각 특징추출기의 가중치 갱신 활성화 여부에 따른 테스트 데이터 세트에 대한 일반적인 합성곱 신경망 모델의 백혈구 세부 유형 분류 성능 지표이다.

Table 3은 분류기를 제외한 합성곱 신경망의 가중치 갱신을 비활성화하여 백혈구 세부 유형 분류 학습을 진행한 일반적인 합성곱 신경망 모델의 성능 지표이며, Table 4는 분류기를 포함하여 합성곱 신경망의 가중치 갱신을 활성화하여 백혈구 세부 유형 분류 학습을 진행하여 혼동행렬을 기반으로 도출한 일반적인 합성곱 신경망 모델의 성능 지표이다. F1-Score를 기준으로 Table 3과 Table 4를 비교해 보면 특징추출기의 가중치 갱신을 활성화하여 학습을 진행한 모델의 성능이 가중치 갱신을 비활성화하여 학습을 진행한 모델보다 더 강인한 성능을 가지는 것을 확인할 수 있다. 따라서, 본 논문에서는 백혈구 세부 유형 분류에 가중치 갱신 활성화를 활용하여 학습을 진행하였다.

Figure 3의 특징 앙상블 단계를 위한 Top-3 합성곱 신경망 모델 선정은 Table 4에서 확인할 수 있는 16개의 합성곱 신경망 모델 중 정확도가 가장 높은 순으로 ConvNeXt Large, ConvNeXt Small, ResNet-152를 선정하였다. Figure 5는 테스트 데이터 세트에 대한 Top-3 합성곱 신경망 모델의 혼동행렬이다.

혼동행렬 분석 결과는 ConvNeXt Large, ConvNeXt Small, ResNet-152 각각 13개, 15개 16개의 오답지를 확인할 수 있다. 호중구 유형에 해당하는 각 모델의 혼동행렬 결과를 확인해 보면 ConvNeXt 기반 모델이 ResNet-152보다 많은 오답지 결과를 확인할 수 있다. 또한, 호산구 유형에 해당하는 결과를 확인해 보면 호중구 유형과는 반대로 ConvNeXt 기반 모델이 ResNet-152보다 강인한 결과를 가진다. 이는 ResNet-152가 ConvNeXt 기반 모델보다 호중구 유형 분류에 유리한 특징을 추출함을 의미하며 ConvNeXt 기반 모델은 호산구 유형 분류에 더 유리한 특징을 추출함을 의미한다. 림프구와 단핵구에 대해서는 모든 모델이 강인한 결과를 가진다. ResNet-152가 추출하는 호중구 분류에 유리한 특징과 ConvNeXt 기반 모델이 추출하는 호산구 분류에 유리한 특징을 특징 앙상블 기법을 통해 모두 활용하여 모든 백혈구 세부 유형 분류에 강인한 성능을 가지는 분류 모델을 설계할 수 있다.

2.4 Classification Stage

Figure 3의 분류 단계에서 제안되는 분류기의 백혈구 세부 유형 분류 성능을 비교 분석하기 위해, Top-3 합성곱 신경망 모델이 출력한 특징 맵을 입력으로 받는 여러 개의 완전 연결 계층(Fully Connected Layer, FC Layer)으로 구성된 분류기를 설계하였으며 분류 성능 분

Table 5. Feature Ensemble Model Performance Results

Models	Precision	Recall	Accuracy	F1-Score
Classifier(w./1 FC Layer)	0.9917	0.9917	0.9958	0.9917
Classifier(w./2 FC Layers)	0.9908	0.9908	0.9954	0.9908
Classifier(w./3 FC Layers)	0.9908	0.9908	0.9954	0.9908
Classifier(w./4 FC Layers)	0.9892	0.9892	0.9946	0.9892

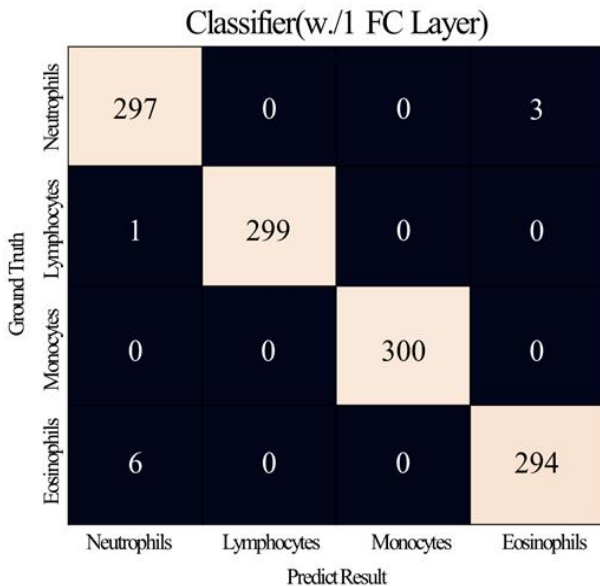


Fig. 6. Classifier(w./1 FC Layer) Confusion Matrix Result

석을 통해 가장 적합한 깊이의 백혈구 세부 유형 분류기를 제안한다. 분류기 설계는 1, 2, 3, 4개의 FC Layer로 구성하여 총 4개의 분류기에 대해 실험을 진행하였으며, 학습 중 과적합을 방지하기 위해 50% 확률로 뉴런을 비활성화하는 Dropout을 적용하였다. Table 5는 FC Layer 깊이에 따른 분류기의 테스트 데이터 세트에 대한 백혈구 세부 유형 분류 평가 결과이다. Table 5에서 확인할 수 있는 것처럼 모든 특징 앙상블 모델이 Precision, Recall, Accuracy, F1-Score에서 98% 이상의 성능을 가지며 Top-3 모델보다 백혈구 세부 유형 분류에 강인한 성능을 가지는 것을 확인할 수 있다. 또한, Figure 6은 1개의 층을 가지는 분류기의 혼동행렬 결과이다. Top-3 모델 중에서 가장 높은 성능을 가지는 Figure 5의 ConvNeXt Large의 혼동행렬 결과와 비교해 보면 호중구를 제외한 유형의 분류 성능 저하가 없음을 확인할 수 있다. 또한, 기존에 호중구 유형 분류에서 가장 강인한 성능을 가지는 ResNet-152보다 높은 분류 성능에 도달하였다.

V. Conclusions and Future Works

본 논문에서는 혈액 현미경 이미지 내 백혈구 세부 유형 분류를 위한 모델을 제안하였다. 이 모델은 백혈구 관심 영역 추출, 특징 앙상블, 분류 세 단계를 통해 높은 성능을 보였다. 혈액 이미지 내 백혈구 탐지 성능 비교 실험을 통해 백혈구 관심 영역 추출 단계를 위한 객체 탐지 알고리즘으로는 백혈구에 대한 객체 탐지 성능이 가장 우수한 YOLO V8s를 선정하였다. 또한, 특징 앙상블 단계에서는 YOLO V8s를 통해 추출한 백혈구 관심 영역에 대한 일반적인 합성곱 신경망의 백혈구 세부 유형 분류 성능 비교 실험을 통해 정확도, 정밀도, 재현율 등 다양한 관점으로 성능을 비교하여 가장 성능이 높은 Top-3 모델인 ConvNeXt Large, ConvNeXt Small, ResNet-152를 선정하였다. 분류 단계에서는 Top-3 모델이 추출한 특징 맵을 기반으로 특징 앙상블을 적용하여 1개의 층으로 설계된 백혈구 세부 유형 분류 기법을 제안하였다. 기존 일반적인 합성곱 신경망 모델보다 백혈구 세부 유형 분류에서 강인한 성능을 정성적, 정량적 평가를 통해 입증하였다. 백혈구는 각각 호중구, 림프구, 단핵구, 호산구 총 4가지 유형을 가지며, 각 유형은 신체 내에 발생한 이상 증세에 따라 해당 증세와 관련된 유형의 백혈구가 증가하는 특징을 가진다. 이러한 특징은 적은 양의 혈액을 통해 환자의 건강 상태를 예측할 수 있기에 지속적인 건강 관리와 질병 예측에 활용 가능한 참고 지표로서 의의를 가진다.

혈액 현미경 이미지 내에는 백혈구뿐만 아니라 적혈구, 혈소판과 같이 다양한 혈액 구성요소가 포함된다. 이는 백혈구 세부 유형 분류에 있어 불필요한 특징이 될 수 있기에, 객체 탐지 기반 알고리즘을 통해 혈액 현미경 이미지 내에서 백혈구에 관한 관심 영역을 추출하였다. 이후, 추출된 관심 영역에 대해 일반적인 합성곱 신경망의 백혈구 세부 유형 분류 실험을 진행하여 성능을 정확도, 정밀도, 재현율과 같이 다양한 측면에서 비교 분석하였다. 해당 결과를 바탕으로 백혈구에 대한 특징 맵을 가장 잘 추출하는 ConvNeXt Large, ConvNeXt Small, ResNet-152로 구성된 Top-3 모델을 선정하였다. 마지막으로, Top-3 모델이 추출한 모든 특징 맵을 입력으로 사용하는 1개의 FC

Layer로 설계된 분류기를 통한 특징 앙상블 모델을 통해 일반적인 합성곱 신경망 모델보다 강한 백혈구 세부 유형 분류 성능을 가진 모델을 제안하였다.

향후 연구에서는 백혈구와 같은 세포 객체의 탐지 및 이상 분류 연구를 통해 신체 내 다양한 세포의 세부 유형 분류 및 이상 탐지 시스템으로 확장하는 것을 목표로 한다. 이를 통해 더욱 정밀한 건강 모니터링 및 질병 예측 시스템을 개발할 예정이다. 본 논문에서 제안한 특징 앙상블 기법에 지식 증류(Knowledge Distillation)와 같은 다양한 학습 기법을 적용하여 백혈구 세부 유형 예측 정확도와 연산 효율성을 개선할 계획이다. 또한, 적은 데이터 세트에서도 강한 성능을 보이는 Data-Efficient Learning 기법을 적용하여 다양한 의료 및 영상 이미지에 대한 인공지능 기술의 적용 가능성을 지속해서 연구할 예정이다.

ACKNOWLEDGEMENT

This research was supported by the Regional Innovation Strategy (RIS) program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (2021RIS-001(1345370811)), and by the Basic Science Research Program through the NRF funded by the Ministry of Education (RS-2023-00248444).

REFERENCES

- [1] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Machine Learning and Knowledge Extraction*, Vol. 5, No. 4, pp. 1680-1716, November 2023. DOI: 10.3390/make5040083
- [2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, Las Vegas, Nevada, USA, June 2016. DOI: 10.1109/CVPR.2016.90
- [3] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700-4708, Honolulu, Hawaii, USA, July 2017. DOI: 10.1109/CVPR.2017.243
- [4] Z. Liu, H. Mao, C. Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11976-11986, New Orleans, Louisiana, USA, June 2022. DOI: 10.1109/CVPR52688.2022.01167
- [5] M. C. Su, C. Y. Cheng, and P. C. Wang, "A Neural-Network-Based Approach to White Blood Cell Classification," *The Scientific World Journal*, Vol. 2014, No. 1, pp. 796371, January 2014. DOI: 10.1155/2014/796371
- [6] O. Saidani, M. Umer, N. Alturki, A. Alshardan, M. Kiran, S. Alsubai, T. H. Kim, and I. Ashraf, "White blood cells classification using multi-fold pre-processing and optimized CNN model," *Scientific Reports*, Vol. 14, No. 1, pp. 3570, February 2024. DOI: 10.1038/s41598-024-52880-0
- [7] T. Wollmann, and K. Rohr, "Deep Consensus Network: Aggregating predictions to improve object detection in microscopy images," *Medical Image Analysis*, Vol. 70, pp. 102019, May 2021. DOI: 10.1016/j.media.2021.102019
- [8] M. L. Ali, Z. Rauf, A. Khan, A. Sohail, R. Ullah, and J. Gwak, "CB-HVT Net: A Channel-Boosted Hybrid Vision Transformer Network for Lymphocyte Detection in Histopathological Images," *IEEE Access*, Vol. 11, pp. 115740-115750, October 2023. DOI: 10.1109/ACCESS.2023.3324383
- [9] X. Zheng, P. Tang, L. Ai, D. Liu, Y. Zhang, and B. Wang, "White blood cell detection using saliency detection and CenterNet: A two-stage approach," *Journal of biophotonics*, Vol. 16, No. 3, e202200174, March 2023. DOI: 10.1002/jbio.202200174
- [10] A. Abdurrazzaq, A. K. bin Junoh, and I. P. R. Yulianto, "White blood cell segmentation by using vector color approach," *International Journal of Applied Mathematics, Sciences, and Technology for National Defense*, Vol. 1, No. 1, pp. 1-6, January 2023. DOI: 10.58524/app.sci.def.v1i1.158
- [11] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, Vol. 9351, pp. 234-241, Munich, Germany, November 2015. DOI: 10.1007/978-3-319-24574-4_28
- [12] Y. Guo, A. I. Shahin, and H. Garg, "An indeterminacy fusion of encoder-decoder network based on neutrosophic set for white blood cells segmentation," *Expert Systems with Applications*, Vol. 246, pp. 123156, July 2024. DOI: 10.1016/j.eswa.2024.123156
- [13] Y. Wu and K. He, "Group Normalization," *Proceedings of the European Conference on Computer Vision*, pp. 3-19, Munich, Germany, September 2018. DOI: 10.1007/978-3-030-01261-8_1
- [14] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common Objects in Context," *European Conference on Computer Vision, Lecture*

- Notes in Computer Science, Vol. 8693, pp. 740–755, Zurich, Switzerland, September 2014. DOI: 10.1007/978-3-319-10602-1_48
- [15] P. Mooney, Blood Cell Images, <https://www.kaggle.com/datasets/paultimothymooney/blood-cells>
- [16] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, Vol. 115, pp. 211–252, April 2015. DOI: 10.1007/s11263-015-0816-y
- [17] I. Radosavovic, R. P. Kosaraju, R. Girshick, K. He, and P. Dollár, "Designing Network Design Spaces," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10425–10433, Seattle, Washington, USA, June 2020. DOI: 10.1109/CVPR42600.2020.01044
- [18] R. Walsh, M. H. Abdelpakey, M. S. Shehata, and M. M. Mohamed, "Automated human cell classification in sparse datasets using few-shot learning," *Scientific Reports*, Vol. 12, No. 1, pp. 2924, February 2022. DOI: 10.1038/s41598-022-06718-2
- [19] C. Y. Wang, I. H. Yeh, and H. Y. Mark Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," *European Conference on Computer Vision*, Lecture Notes in Computer Science, pp. 1–21, Mico Milano, Italy, September 2024. DOI: 10.1007/978-3-031-72751-1_1
- [20] S. Aharon, Louis-Dupont, O. Masad, K. Yurkova, F. Lotem, Lkdci E. Khvedchenya, R. Rubin, N. Bagrov, B. Tymchenko, T. Keren, A. Zhilko and Eran-Deci, "Super-Gradients," GitHub repository, 2021, DOI: 10.5281/ZENODO.7789328
- [21] S. Ahmad, Z. Ullah, and J. Gwak, "Multi-teacher cross-modal distillation with cooperative deep supervision fusion learning for unimodal segmentation," *Knowledge-Based Systems*, Vol. 297, pp. 111854, August 2024. DOI: 10.1016/j.knosys.2024.111854
- [22] M. U. Farooq, Z. Ullah, A. Khan, and J. Gwak, "DC-AAE: Dual channel adversarial autoencoder with multitask learning for KL-grade classification in knee radiographs," *Computers in Biology and Medicine*, Vol. 167, pp. 107570, December 2023. DOI: 10.1016/j.combiomed.2023.107570
- [23] Z. Rauf, A. R. Khan, A. Sohail, H. Alquhayz, J. Gwak, and A. Khan, "Lymphocyte detection for cancer analysis using a novel fusion block based channel boosted CNN," *Scientific Reports*, Vol. 13, No. 1, pp. 14047, August 2023. DOI: 10.1038/s41598-023-40581-z

Authors



Junhwi Park received the B.S. degree in AI-Robotics Engineering from Korea National University of Transportation (KNUT) in 2024. From 2024, he has been working as a Master's student and researcher of Algorithmic Machine

Intelligence Laboratory in KNUT. His current research interests include deep learning, computer vision, image and video processing.



Changjoon Park received the B.S. degree in AI-Robotics Engineering from Korea National University of Transportation (KNUT) in 2023. From 2023, he has been working as a Master's student and researcher in the Algorithmic

Machine Intelligence Laboratory in KNUT. His current research interests include machine learning/deep learning.



Namjung Kim received a B.S. degree in Software from Korea National University of Transportation (KNUT) in 2023. From 2023, he has been working as a Master's student and researcher in the Algorithmic Machine

Intelligence Laboratory in KNUT. He has been studying Time-Series Analysis and Multimodal Analysis in deep learning and machine learning.



Ryumduck Oh is a Professor in the department of Software at Korea National University of Transportation. He received the Ph.D. in Computer Science at Hongik University Seoul, Korea, in 1993. He received the B.S. and M.S.

degrees in Computer Science from Hongik University in 1986 and 1988, respectively. His research interest includes the database, big data analysis, and the IoT platform on railway environment.



Jeonghwan Gwak received the Ph.D. degree in machine learning and artificial intelligence from Gwangju Institute of Science and Technology (GIST), Gwangju, Korea in 2014. From 2002 to 2007, he worked for

several companies and research institutes as a researcher and a chief technician. From 2014 to 2016, he worked as a postdoctoral researcher at GIST, and from 2016 to 2017 as a research professor. From 2017 to 2019, he was a research professor in Biomedical Research Institute & Department of Radiology at Seoul National University Hospital, Seoul, Korea. In 2019, he joined Korea National University of Transportation (KNUT) as an Assistant Professor and since 2024, he has been a Professor. He is the director of Algorithmic Machine Intelligence Laboratory. His current research interests include deep learning, computer vision, image and video processing, AIoT (Artificial Intelligence of Things), fuzzy modeling, evolutionary algorithms, optimization, machine unlearning, and applications in medical imaging, smart healthcare systems, energy systems, transportation, and visual surveillance.