

자연스러운 이미지 가명화를 위한 Mask R-CNN과 SRNet 기반 개인정보 보호 시스템 구현

한영욱*, 우혜진**, 우지호***, 김규리****, 이수정*****, 유길상*****

Implementation of a Personal Information Protection System based on Mask R-CNN and SRNet for Natural Image Pseudonymization

Younguk Han*, Hyejin Woo**, Jiho Woo***, Kyuri Kim****, Sujung Lee*****, and Gilsang Yoo*****

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(RS-2023-00246191)

요 약

소셜미디어의 활성화로 공개된 사진에 대한 개인정보 유출이 증가하고 있다. 개인정보 유출을 방지하기 위한 현재 사용되는 비식별화 기법은 개인정보에 해당하는 영상을 인위적으로 왜곡시켜 데이터의 활용도를 저하시키고 이미지의 자연스러움을 훼손하는 한계를 가진다. 본 연구는 이미지 데이터의 개인정보 객체를 효과적으로 탐지하고 개인정보 영역을 자연스럽게 가명화하는 시스템을 제안하였다. Mask R-CNN 기반의 객체 탐지 모델을 활용하여 신분증과 간판 등 개인정보 객체를 식별하고, SRNet을 적용하여 해당 정보를 자연스러운 가명 텍스트로 대체하였다. 실험 결과 신분증 탐지 모델은 평균 정밀도(mAP)는 0.947을 기록하여 높은 탐지 성능을 보였으며, SRNet 기반 가명화 기법은 이미지의 시각적 자연스러움을 유지하며 개인정보 보호 효과를 극대화하였다. 제안 시스템은 향후 비정형 데이터의 개인정보 보호 및 안전한 활용에 실무적으로 기여할 것으로 기대된다.

Abstract

With the proliferation of social media, the leakage of personal information through publicly shared images has increased significantly. Current anonymization techniques designed to prevent such leakage often involve artificially distorting visual data containing personal information, which reduces data utility and compromises image naturalness. This study proposes a system that effectively detects personal information objects in unstructured image data and seamlessly pseudonymizes the corresponding regions. A Mask R-CNN-based object detection model was employed to identify personal information objects such as identification cards and signboards, and SRNet was applied to replace the detected content with visually natural pseudonymous text. Experimental results demonstrate that the ID card detection model achieved a mean Average Precision (mAP) of 0.947, indicating high detection performance. Moreover, the SRNet-based pseudonymization technique preserved the visual naturalness of images while enhancing personal information protection. The proposed system is expected to contribute practically to the protection and secure utilization of personal information in unstructured data.

Keywords

Mask R-CNN, instance segmentation, scene text recognition, SRNet, pseudonymization

* 숭실대학교 경영학부
- ORCID: <https://orcid.org/0009-0008-8204-1231>
** 고려대학교 인공지능사이버보안학과
- ORCID: <https://orcid.org/0009-0009-9151-603X>
*** 고려대학교 사범대학 체육교육과
- ORCID: <https://orcid.org/0009-0001-5263-9936>
**** 중앙대학교 소프트웨어학부
- ORCID: <https://orcid.org/0009-0004-7543-1945>
***** 고려대학교 컴퓨터융합소프트웨어학과
- ORCID: <https://orcid.org/0009-0000-0362-9996>

***** 고려대학교 정보대학 정보창의교육연구소(교신저자)
- ORCID: <https://orcid.org/0009-0002-1085-5355>

• Received: Apr. 07, 2025, Revised: Jun. 10, 2025, Accepted: Jun. 13, 2025
• Corresponding Author: Gilsang Yoo
Creative Informatics & Computing Institute, 145 Anam-ro,
Seongbuk-gu, Seoul, S.Korea
Tel.: +82-2-3290-1674, Email: ksyoo@korea.ac.kr

1. 서 론

최근 소셜미디어와 디지털 기술이 급속히 발전하면서 개인이 온라인에 게시한 이미지 데이터가 무단으로 수집·활용되는 사례가 급증하고 있다. 특히 개인이 일상적으로 촬영하여 공유하는 이미지 내에는 주민등록번호, 이름, 주소, 전화번호 등 민감한 개인정보가 본인의 의사와 무관하게 포함될 수 있으며, 이러한 정보가 제3자에 의해 무단으로 수집·유통될 경우 심각한 개인정보 침해로 이어질 수 있다.

개인정보보호위원회에 따르면, 2024년 한 해 동안 개인정보 침해 상담 및 신고 건수가 지속적으로 높은 수준을 유지하고 있으며, 특히 '적법하지 않은 개인정보 이용 또는 제공'과 '영상정보처리기기 설치·운영 제한' 분야에서 높은 발생 빈도를 기록하고 있다[1]. 대한민국 정부에서는 2016년 6월에 개인정보 비식별 조치 가이드라인을 제시하였다. 개인정보 비식별 조치란 '개인정보에서 개인을 식별할 수 있는 요소를 전부 또는 일부 삭제하거나 대체하는 등의 방법을 활용하여 개인을 알아볼 수 없도록 하는 조치'를 말한다. 또한 개인정보 보호법에서는 비식별 조치의 방식으로 가명처리와 익명처리를 인정하고 있다. 가명처리란 '개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것'으로 정의된다(개인정보 보호법 제2조 제1호의2). 2024년 개인정보보호위원회가 발표한 가이드라인에서는 기업 및 기관을 대상으로 마스킹(Masking)과 블러링(Blurring) 기술을 활용한 비식별 조치를 권고하고 있다[2]. 이에 따라 현행 개인정보 보호 조치는 주로 기업이나 조직 단위에서 이루어지는 마스킹이나 블러링과 같은 비식별화 기법을 중심으로 하고 있으나, 이러한 방법들은 데이터의 유용성을 크게 저하시키거나 이미지의 자연스러움을 훼손하는 단점을 가진다. 이에 반해, 가명처리 방식은 정보를 가리는 것이 아니라 대체한다는 점에서 이미지의 자연스러움을 유지한다. 동시에, 모자이크 및 블러링 기법에 반해 복호화가 불가능하다는 장점 또한 지니고 있다. 또한, 소셜미디어 환경에서는 개인의 이미지가 무단으로 사용되거나 유통되는 사

례가 빈번히 발생하고 있으나, 이러한 이미지 내에서 개인정보를 효과적으로 식별하고 제거하기 위한 기술적 방안이나 명확한 가이드라인이 부족하다. 특히 신분증이나 간판 등과 같이 소셜미디어에 의도치 않게 노출되는 개인정보 및 간접식별정보는 단순 가림 처리만으로는 오히려 시각적 이질감을 유발해 이미지의 진정성이나 공유 의도를 훼손시킬 수 있다. 이는 사용자들이 콘텐츠를 자율적으로 관리하거나 통제하는 데 어려움을 초래하며, 결과적으로 개인정보 보호에 실질적인 한계를 발생시킨다. 따라서 시각적 자연스러움을 유지하면서도 식별이 불가능하도록 처리할 수 있는 정교한 기술적 접근이 필요하다.

이러한 문제의식을 바탕으로, 본 연구에서는 이미지 데이터 내 개인정보 객체를 자동으로 탐지하고, 원본 이미지의 훼손 없이 시각적 자연스러움을 유지하면서도 식별이 불가능하도록 가명화하는 시스템을 제안하였다. 비정형 이미지 데이터 내 개인정보 객체를 효과적으로 탐지하기 위해 딥러닝 기반의 Mask R-CNN(Mask Region-based Convolutional Neural Networks) 객체 탐지 모델을 활용하였으며, 탐지된 영역을 시각적으로 자연스럽게 가명화하기 위해 SRNet(Style Retention Network) 기반의 텍스트 변환 기법을 적용한 가명처리 시스템을 제안하고 구현하였다. 제안된 시스템은 신분증, 간판 등 개인정보가 포함된 대표적인 객체를 정확하게 탐지하고 자연스럽게 실용적인 방식으로 가명화함으로써 기존 비식별화 기법의 한계를 극복하고자 한다. 이를 통해 개인 및 조직 차원의 개인정보 보호 효과를 강화하고, 비정형 데이터의 보다 안전한 활용 가능성을 제시한다.

본 논문의 구성은 다음과 같다. 제2장에서는 가명처리와 관련된 기존 연구를 고찰하고, 객체 탐지 및 가명처리 기술의 이론적 배경을 소개한다. 제3장에서는 본 연구에서 제안하는 비정형 데이터 가명처리 시스템의 설계 및 구현 방법을 기술하였다. 특히 Mask R-CNN를 이용한 인스턴스 세그멘테이션(Instance segmentation) 기반 객체 탐지 과정과 EasyOCR(Optical Character Recognition) 및 CRAFT(Character Region Awareness for Text Detection)를 활

용한 텍스트 영역 인식 과정, 그리고 SRNet를 이용한 자연스러운 텍스트 변환 기법을 기술하였다.

제4장에서는 제안된 시스템의 성능을 평가하기 위한 실험 환경과 방법을 기술하고, 객체 탐지 및 가명화 성능의 정확도를 분석하여 결과를 비교한다. 마지막으로 제5장에서는 본 연구의 결론과 향후 연구 방향에 대해 기술하였다.

II. 관련 연구

2.1 가명처리

현대 사회에서 개인정보는 기업의 핵심 자산이자 보호가 필요한 민감 정보로 분류되며, 유출 시 개인의 권리를 심각하게 침해할 수 있기 때문에 철저한 보호가 요구된다. 그러나 과도한 보호는 데이터 기반의 기술 혁신을 저해할 수 있는 한계가 있으며, 이러한 문제를 해결하기 위해 도입된 가명정보는 개인정보를 식별할 수 없도록 처리함으로써 보호와 활용의 균형을 가능하게 한다. 이미지 및 영상 분야에서는 주로 얼굴이나 신체 특성을 가리는 방식으로 비식별화 기술이 연구되어 왔으며, M. Maximov et al.[3]은 사람의 얼굴 및 신체 특징을 가명처리하는 방법을, H. Hukkelåset al.[4]는 이미지 상의 얼굴 가명처리 기법을 제안하였다. 이 외에도 의료 데이터 분야에서의 가명정보 활용 연구가 활발히 진행되고 있다[5]. 하지만 코로나19 팬데믹 이후 비대면 활동이 증가하면서, 소셜 이미지에 포함된 신분증, 택배송장 등 다양한 형태의 민감 정보로 인한 개인정보 유출 문제가 새롭게 대두되었으며, 이에 따라 보다 포괄적이고 정교한 가명처리 기술에 대한 연구의 필요성이 커지고 있다[6][7].

2.2 Instance segmentation 기술

Mask R-CNN은 객체 탐지(Object detection)와 인스턴스 세그멘테이션을 동시에 수행할 수 있도록 설계된 딥러닝 기반 모델이다[8]. 이 모델은 Faster R-CNN을 기반으로 하며, 기존의 경계 상자(Bounding box) 예측에 더해 객체의 픽셀 단위 마스크

크까지 예측할 수 있도록 구조가 확장되었다. Faster R-CNN은 두 단계로 구성된 검출 방식을 통해 높은 탐지 정확도를 제공하지만, 픽셀 수준의 정밀한 분할에는 한계가 존재한다[9]. 이러한 한계를 보완하기 위해 Mask R-CNN은 기존 구조에 객체별 마스크를 생성하는 분할 헤드(Segmentation head)를 추가하였다. 또한, ROI(Region of Interest) 영역에서 보다 정확한 공간 정렬을 가능하게 하는 RoIAlign 연산을 도입함으로써, 기존 RoIPooling에서 발생하는 위치 불일치 문제를 해결하고 객체의 경계 정보를 더욱 정밀하게 학습할 수 있도록 하였다.

Mask R-CNN은 객체 분류, 경계 상자 회귀, 마스크 예측의 세 가지 작업을 병렬적으로 수행하며, 이를 다중 손실 함수(Multi-task loss function)를 통해 동시에 최적화한다. 해당 다중 손실 함수는 식 (1)과 같이 정의된다.

$$L = L_{cls} + L_{box} + L_{mask} \quad (1)$$

여기서 L_{cls} 는 객체 분류 손실, L_{box} 는 경계 상자 회귀 손실, L_{mask} 는 마스크 예측 손실을 의미하며, 각각의 손실을 독립적으로 최적화하여 모델의 성능이 향상된다. 기본적으로 ResNet-50 또는 ResNet-101과 같은 백본을 사용하여 입력 이미지에서 특징을 추출하며, FPN(Feature Pyramid Network)과 결합하여 다양한 크기의 객체를 효과적으로 탐지할 수 있도록 한다[10].

그림 1은 Mask R-CNN의 전체적인 구조를 보여준다. Faster R-CNN의 두 가지 주요 모듈인 RPN(Region Proposal Network)과 RoI Head에 추가적으로 마스크 분할 헤드가 포함되어 있으며, RoIAlign을 통해 보다 정밀한 특징 정렬이 이루어진다. 이를 통해 Mask R-CNN은 COCO(Common Objects in Context) 데이터셋과 같은 대규모 벤치마크에서 뛰어난 성능을 보이며, 의료 영상 분석, 자동 운전, 증강 현실 등의 다양한 응용 분야에서 활용될 수 있다. 또한, 객체 탐지뿐만 아니라 사람 자세 추정(Human pose estimation)과 같은 다른 비전 문제에도 확장하여 사용할 수 있다.

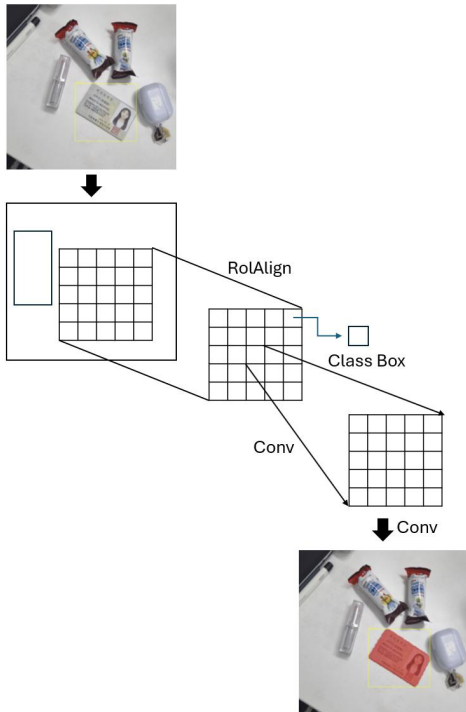


그림 1. instance segmentation을 위한 Mask R-CNN 프레임 워크

Fig. 1. Mask R-CNN framework for instance segmentation

2.3 Scene text detection/recognition 기술

CRAFT는 이미지 내 텍스트 영역을 추출하기 위해 설계된 딥러닝 기반의 텍스트 검출 모델이다 [11]. 특히 다양한 글꼴, 크기, 배경을 가진 장면 텍스트에 대해 강인한 성능을 보이며, 단어 단위가 아닌 문자 단위로 영역을 탐지한다. 각 문자의 Region Score와 Affinity Score를 예측하여 문자를 단어 단위로 연결하는 방식으로, 복잡한 배경이나 왜곡된 텍스트도 효과적으로 인식할 수 있다[12]. 본 연구에서는 일상적인 이미지 내의 다양한 형태의 텍스트 탐지를 위해 CRAFT 모델을 텍스트 검출 단계에 활용하였다.

EasyOCR은 텍스트 인식을 위한 딥러닝 기반 광학 문자 인식 프레임워크로, 80개 이상의 언어를 지원하며 이미지에서 텍스트의 위치와 내용을 동시에 추출할 수 있다[13][14]. EasyOCR은 CRAFT를 텍스트 검출에, CRNN(Convolutional Recurrent Neural Network)을 텍스트 인식에 사용하여 높은 정확도의 문자 인식을 제공한다[15]. 또한 PyTorch 기반으로 개발되어 다른 모델과의 통합 및 확장이 용이하다.

본 연구에서는 주민등록증 내 다양한 항목을 분류하고, 해당 카테고리에 맞는 가명 정보를 삽입하기 위해 한국어 인식 성능이 우수한 EasyOCR을 적용하였다.

2.4 Scene text editing 기술

본 연구의 핵심 기술은 가명 처리를 수행하더라도 원본 이미지와 비교했을 때 시각적으로 ‘어색함이 없는’ 자연스러운 이미지를 생성하는 데 있다. 즉, 기존의 개인정보 텍스트를 자연스럽게 대체할 수 있는 가상 데이터를 생성하여 원본 배경과 조화롭게 합성하는 것이다. 이를 위해 다양한 이미지 생성 및 변환 기술의 적용 가능성을 검토하였다.

우선, GAN(Generative Adversarial Network) 기반의 모델들인 Cycle-GAN과 Pix2Pix는 가명처리 영역에 활용 가능한 대표적인 이미지-이미지 변환 모델로 고려되었다. Cycle-GAN은 paired data 없이 도메인 간 이미지 변환이 가능하다는 장점을 가지고 있으며, 텍스트와 그에 대응하는 이미지 쌍을 별도로 생성하지 않아도 학습이 가능하다는 점에서 효율적인 학습 구조를 제시한다[16]. Pix2Pix는 이미지 내 시각적 요소, 특히 배경과 구성을 자연스럽게 변환하는 데 강점을 보여, 텍스트 영역을 포함한 이미지 변환에 적용 가능성이 평가되었다[17]. 그러나 이들 모델은 주로 이미지 스타일 전송에 최적화되어 있으며, 텍스트의 의미를 보존하며 직접적으로 편집하는 데는 한계가 있다. 특히, 텍스트의 결손을 보완하기 위한 이미지 복원 기법으로 Telea 또는 Navier-Stokes 알고리즘 기반의 전통적인 인페인팅 방식이 더 자연스러운 결과를 제공한다[18].

또한, Text Style Transfer 모델의 경우 대부분 영어 또는 라틴 문자 기반의 텍스트에 초점을 맞추고 있어, 한국어와 같은 비라틴 문자에 대한 처리를 위해서는 별도의 데이터 학습이 필요하다는 제약이 있다. 이러한 한계를 보완하기 위해 본 연구에서는 SRNet을 채택하였다. SRNet은 장면 텍스트 편집에 특화된 모델로, 텍스트 내용은 변경하면서도 원본 이미지의 배경 스타일, 조명, 그림자, 기물기 등 시각적 특징을 유지할 수 있는 강력한 성능을 갖춘 모델이다[19].

이를 통해 텍스트와 배경 간의 일체감을 유지하면서도 현실적인 수준의 자연스러운 이미지 생성이 가능하다. SRNet을 활용한 텍스트 편집 및 시물레이션 결과는 그림 2와 같다.

또 다른 Scene Text Editing 모델로는 STEFANN (Scene Text Editor using Font Adaptive Neural Network)이 있으며, 이는 다양한 폰트 특성에 적응하여 텍스트를 편집하는 방식으로 동작한다[20]. 그러나 한글 대체 텍스트를 적용한 실험 결과, 그림 3에서와 같이 SRNet에 비해 시각적 품질이 현저히 떨어지는 성능을 보였다. 또한 구조적 유사도(SSIM, Structural Similarity Index)를 기준으로 성능을 비교한 결과, SRNet이 0.7900, STEFANN이 0.7712로 나타나 SRNet이 보다 우수한 이미지 자연스러움을 유지함을 확인할 수 있다[8][9].



그림 2. SRNet의 Scene text editing 예시
Fig. 2. Example of scene text editing of SRNet



그림 3. STEFANN 테스트 예시
Fig. 3. STEFANN test result

SRNet은 세 가지 주요 모듈로 구성되며, 각 모듈의 구조는 그림 4와 같다. TCM(Text Conversion Module)은 원본 텍스트를 새로운 텍스트로 변환하는 역할을 하며, Stroke Skeleton 알고리즘을 기반으로 텍스트의 의미적 구조와 스타일을 최대한 유지하면서 현실감 있는 전경 텍스트를 생성한다. BIM(Background Inpainting Module)은 텍스트가 위치한 영역의 픽셀을 제거한 후, 하향식 특징 융합(Top-down feature fusion) 방식을 적용하여 해당 영역을 자연스러운 질감으로 복원한다. 이 과정에서는

U-Net 아키텍처가 채택되어 배경의 세부 질감을 효과적으로 복원한다. Fusion Module은 앞선 두 모듈에서 생성된 전경 텍스트 정보와 배경 텍스트 정보를 통합하여 최종적으로 자연스럽고 일관된 편집 텍스트 이미지를 생성한다. 다만, 한국어 데이터베이스를 기반으로 학습되지 않아 한글 텍스트에 대한 처리 성능은 상대적으로 낮다는 한계가 존재한다.

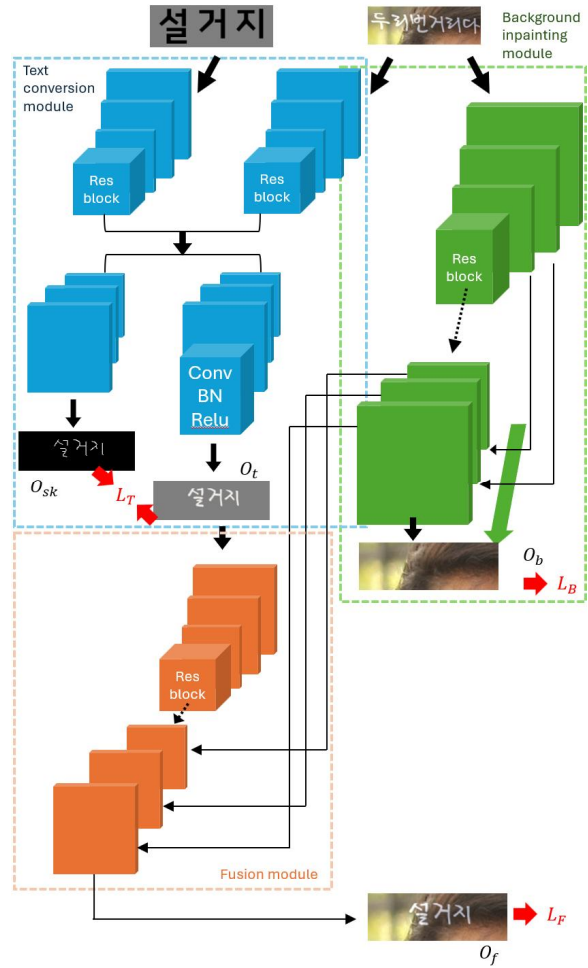


그림 4. SRNet 전체 구조도
Fig. 4. Overall structure of SRNet

III. 제안한 객체 탐지 및 가명 처리 모델

제안한 가명처리 모델은 일상 사진을 입력으로 받아, 개인정보 유출 가능성이 높은 신분증 및 간판 객체를 대상으로 가명화를 수행하도록 구현된다. 시스템은 입력 이미지로부터 개인정보 객체를 자동으로 탐지하고, 탐지된 영역에 대해 자연스러운 가명화를 적용한다. 전체 시스템의 구조는 그림 5와 같다.

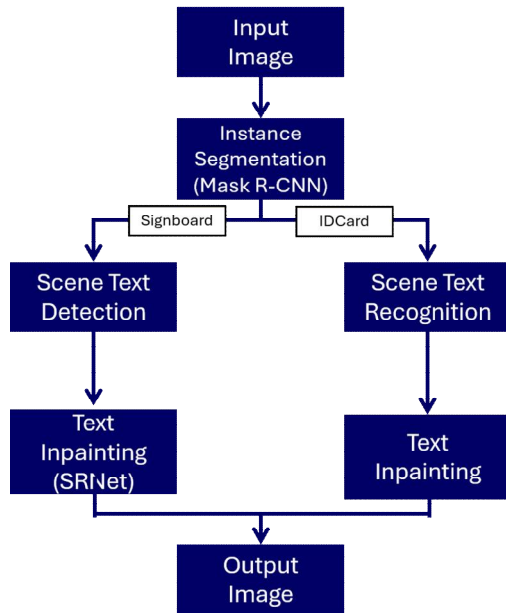


그림 5. 전체 시스템 구조도
Fig. 5. Overall system architecture

우선, 사용자가 입력한 이미지는 Mask R-CNN을 기반으로 분석되며, 인스턴스 세분화를 통해 개인정보 유출 가능성이 있는 객체인 신분증 및 간판을 탐지한다. 탐지된 객체의 종류에 따라 각각 다른 가명처리 프로세스를 수행한다.

신분증 객체의 경우, 고정된 크기로 정규화된 후 선형 변환(Linear transformation) 과정을 거친다. 이후, 화면 텍스트 탐지를 통해 객체 내부의 텍스트를 인식하고, 정규표현식을 활용하여 인식된 문자가 주민등록번호, 이름 등 어떤 개인정보 항목에 해당하는지를 판별한다. 이에 따라 해당 위치에 적절한 가짜 정보를 삽입하여 신분증에 대한 가명처리를 수행한다.

간판 객체는 별도의 변환 없이 화면 텍스트 감지를 통해 텍스트가 포함된 영역을 식별하고, 해당 영역에 SRNet 기반의 텍스트 편집(Scene text editing) 기술을 적용하여 자연스럽게 가명화한다.

최종적으로, 각 객체에 대해 수행된 가명처리 결과는 원본 이미지에 자연스럽게 합성되어 사용자에게 결과 이미지 형태로 제공된다.

3.1 데이터 수집 및 데이터 셋 구축

본 연구에서는 다음의 세 가지 데이터셋을 활용

하였다. 첫째, 신분증이 포함된 이미지 8,550장을 직접 수집하고 수작업으로 라벨링하여 학습용 데이터를 구축하였다. 둘째, 간판 데이터는 Roboflow에 공개된 한국어 간판 이미지 2,853장을 수집하여 활용하였다. 셋째, SRNet 모델의 한글 가명화 성능 향상을 위해 다양한 배경, 글자 색상, 형태, 글꼴을 포함한 한글 및 숫자 혼용 텍스트 데이터를 20만 장 생성하였다.

3.2 객체 탐지

객체 탐지 단계에서는 Mask R-CNN 모델을 활용하여 신분증과 간판을 탐지하였다. Mask R-CNN은 다양한 객체 탐지 기능을 모듈화한 프레임워크로, 기본적으로 우수한 탐지 성능을 보이나, 특정 작업에서의 성능 향상을 위해서는 추가적인 학습이 요구된다. 신분증과 간판 객체를 보다 정확히 탐지하기 위해 다음과 같은 조건으로 모델을 학습하였다.

먼저, Mask R-CNN의 인스턴스 분할 성능을 최적화하기 위해 하이퍼파라미터를 조정하였다. 기본 백본인 ResNet-50-FPN 대신 더 깊은 네트워크 구조인 ResNet-101-FPN을 사용하여 복잡한 객체 경계의 학습 성능을 향상시켰다. 또한, RoIAlign 연산을 적용하여 픽셀 정렬 오류를 최소화하고, 마스크 예측 해상도를 기존 28×28에서 56×56으로 상향 조정하여 분할의 정밀도를 높였다.

모델 학습은 COCO데이터셋에서 사전 학습된 가중치를 초기값으로 설정한 후, 이를 고정한 채 마지막 레이어 위주로 전이 학습을 수행하였다. 최적화 과정에서는 SGD(Stochastic Gradient Descent) 옵티마이저를 사용하였으며, 초기 학습률은 0.02, 모멘텀은 0.9, 가중치 감쇠는 0.0001로 설정하였다. 학습률은 스텝 학습률 감소(Step decay) 방식으로 조정하여 특정 에폭(Epoch)마다 점진적으로 감소시켰다. 학습은 총 24 에폭동안 수행되었으며, 배치 크기는 16으로 설정하였다. 에폭은 기계학습과 딥러닝에서 모델을 학습시키는 반복 횟수를 의미하며, 전체 학습 데이터셋이 한 번 모델에 모두 전달되어 학습이 완료되는 것을 한 에폭이라고 한다. 모델의 일반화 성능을 높이기 위해 데이터 증강 기법으로 랜덤 크롭(Crop)과 수평 뒤집기를 적용하였다.

또한 학습, 검증, 실험 데이터셋은 79%, 14%, 7%의 비율로 학습을 수행하였다.

표 1과 같은 하이퍼파라미터 최적화를 통해 다양한 크기의 객체에 대해 높은 분할 정확도와 정밀한 경계 예측 성능을 보였다.

표 1. Mask R-CNN 모델의 하이퍼파라미터 최적값
Table 1. Optimal hyperparameter values for the mask R-CNN model

Optimizer	Stochastic gradient descent
Learning rate	0.02
Momentum	0.9
Weight decay	0.0001
Batch size	16
Epoch	24
train set ratio	79%
validation set ratio	14%
test set ratio	7%

그림 6은 학습된 모델을 테스트 이미지에 적용한 결과로, 유사한 객체(신분증, 신용카드)들 사이에서도 신분증만을 정확하게 탐지한 성능을 확인할 수 있다. 그림 6의 (b)는 개인정보 추출을 위해 촬영된 영상으로부터 정확히 신분증만 탐지하고 텍스트 정보를 추출하는 과정이다.



(a) 객체 검출 단계
(a) Object detection step



(b) 영역 분리 단계
(b) Area separation step
그림 6. 신분증 객체 탐지 결과
Fig. 6. Id card detection output

본 파이프라인은 두 단계로 구성된다: (1) 복잡한 장면에서 정확히 신분증을 탐지하는 객체 검출 단계, (2) 마스크 기반 분할 또는 투시 보정을 통해 문서 영역을 분리하는 단계. 이러한 과정을 통해 본 시스템은 그림 6, 그림 7과 같이 개인정보 유출이 일어날 수 있는 객체를 정확하게 탐지하는 성능을 확보하였다.



그림 7. 간판 객체탐지 결과
Fig. 7. Signboard detection output

3.3 가명화 처리

가명 정보를 신분증 내 적절한 위치에 적용하기 위한 절차는 다음과 같다.

먼저, 신분증 내 문자인식(OCR)의 정확도를 높이기 위한 전처리 단계로 선형 변환 과정을 적용한다. 이 과정에서는 탐지된 신분증 영역의 점 집합 P 에 대해 Convex Hull을 계산하고, 이를 포함하는 최소 면적의 직사각형을 구하여 네 개의 꼭짓점을 추출한다. 이후 해당 네 꼭짓점을 실제 신분증의 규격인 가로 8.6cm, 세로 5.4cm의 비율에 맞추어 선형 변환을 수행하였다. 선형 변환에는 식 (2)에서 정의한 변환 공식을 적용한다.

$$\begin{bmatrix} wx' \\ wy' \\ w \end{bmatrix} = M_P \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

선형 변환 과정을 거친 이미지에 대해 OCR을 적용하여, 그림 8과 같이 신분증 내 텍스트를 인식하고, 이를 후속 단계에서 활용한다.

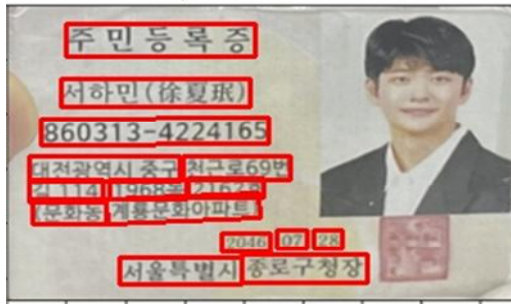


그림 8. 신분증 OCR 탐지 결과
Fig. 8. Id card OCR output

OCR을 통해 추출된 텍스트는 정규표현식과 위치 정보를 기반으로 이름, 주민등록번호, 주소 등으로 분류되며, 해당 항목에 적절한 가명 데이터를 삽입하여 가명처리를 수행한다. 이후 가명 처리된 텍스트를 원본 이미지 상에 자연스럽게 삽입함으로써 신분증의 가명화가 완료된다.

간판의 경우, 내부의 문자를 탐지한 후 해당 영역에 SRNet 기반의 텍스트 변환 기술을 적용하여 가명처리를 수행한다. 이를 위해 Scene Text Detection 단계를 통해 문자 영역을 탐지하며, CRAFT 모델을 활용해 그림 9와 같이 간판 이미지로부터 텍스트 영역을 추출한다.



(a) 원본 사진
(a) Original image



(b) 추출된 글씨 영역
(b) Extracted text region
그림 9. 간판 내 글씨 추출

Fig. 9. Text extracted from signboard

기본 SRNet 모델은 한글 텍스트의 변환 성능이 낮다는 한계가 있으므로 이를 해결하기 위해 별도로 구축한 한글 데이터를 기반으로 모델을 재학습시켜 한글 텍스트도 안정적으로 인식 및 가명화가 가능하다. 그림 10은 한글 학습 전후의 SRNet 결과를 비교한 것으로, 간판의 글씨체를 유지하면서 자연스럽게 가명화가 수행된다.



(a) 추출된 글씨 영역
(a) Extracted text region



(b) 가명화 처리 결과
(b) Pseudonymization result

그림 10. SRNet 학습 적용 후 가명처리 결과
Fig. 10. Pseudonymization results after applying SRNet training

텍스트 편집 모델 학습은 자체 제작 데이터셋에서 표 2와 같은 하이퍼파라미터 최적화를 적용하였다. 초기 학습률을 0.001, 0.9의 지수 감소율을 적용하고 Adam 옵티마이저를 적용하였으며, 베타1은 0.9, 베타2는 0.999를 적용하였다.

표 2. SRNet 모델의 하이퍼파라미터 최적값
Table 2. Optimal hyperparameter values for the SRNet model

Optimizer	Adam
Learning rate	0.001
Beta1	0.9
Beta2	0.999

IV. 실험 결과

4.1 신분증 탐지 모델 학습 및 평가

신분증 탐지 정확도를 높이기 위해 다양한 환경에서 촬영된 신분증 이미지 8,550장을 직접 제작하여 학습 데이터로 활용하였다. 이 데이터는 조명, 각도, 배경 등 다양한 실제 상황을 반영하고 있어, 모델이 보다 현실적이고 복잡한 조건에서도 효과적으로 학습될 수 있다. 학습 과정에서는 에폭의 증가에 따라 인스턴스 세그멘테이션의 예측 범위 및 객체 탐지 성능과 관련된 손실 값이 점차 감소하였으며, 이는 모델이 점진적으로 신분증 객체를 보다 정확하게 인식하고 분류하는 방향으로 학습되었다.

Mask R-CNN 모델을 기반으로 한 신분증 학습 결과는 그림 11과 같다.

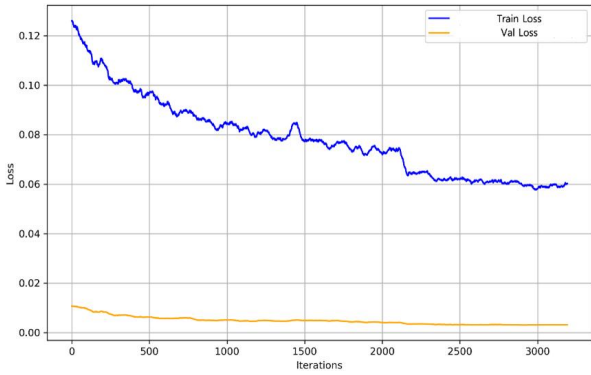


그림 11. 신분증 학습 반복 횟수에 따른 검증 및 검증 손실 변화

Fig. 11. Training and validation loss trends iterations for Id cards

학습 손실 값은 초기 약 0.12에서 시작하여 에폭 24 시점에는 약 0.6까지 감소하는 추세를 나타내었으며, 이는 훈련 데이터에 대한 모델의 성능이 지속적으로 개선되고 있음을 시사한다. 또한, 검증 손실 (Validation Loss) 역시 에폭이 진행됨에 따라 전반적으로 감소하는 경향을 보였으며, 이는 과적합 없이 모델이 일반화 성능을 확보하고 있음을 의미한다.

그림 12와 같이, BBox와 Segmentation의 학습 정확도는 에폭이 증가함에 따라 점진적으로 향상되는 양상을 보였다. BBox의 경우, 초기 mAP는 0.89에서 시작하여 에폭 12에 mAP가 약 0.93에 도달하였으며, 이후에도 지속적인 성능 개선이 나타났다. 세그멘테이션 역시 초기 mAP 0.92에서 시작해 에폭 12 기준 mAP가 약 0.947까지 상승하며 꾸준한 향상세를 보였다. 이러한 결과는 모델이 학습 데이터를 점차 효과적으로 학습하며 예측 성능을 향상시키고 있다.

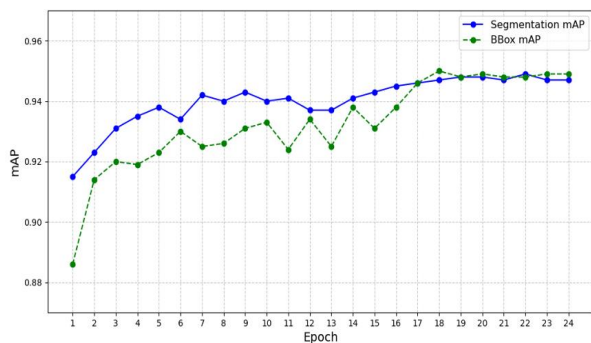


그림 12. 신분증 학습 정확도의 에폭별 변화
Fig. 12. Train accuracy over epochs for Id card

기존의 신분증 인스턴스 세그멘테이션 연구들은 주로 MIDV-500 데이터 셋을 활용하여 정적인 환경에서 단일 객체(한 장의 신분증)를 대상으로 하며, 근거리 촬영된 이미지를 기반으로 학습된 경우가 주를 이룬다[21]. 반면, 본 연구에서는 신분증이 원 거리에 위치하거나 여러 장이 동시에 존재하는 복잡한 장면에서도 효과적으로 탐지 및 분할이 가능하도록 모델을 학습하였으며, 이러한 점에서 기존 연구들과의 차별성을 지닌다.

4.2 간판 탐지 모델 학습 및 평가

간판 탐지 정확도를 향상하기 위해 다양한 환경에서 촬영된 간판 이미지를 학습 데이터로 활용하였다. 이러한 데이터는 실제 상황을 반영하고 있어 모델이 보다 현실적인 조건에서 정교하게 학습될 수 있도록 한다. 학습이 진행됨에 따라 인스턴스 세그멘테이션과 객체 탐지와 관련된 손실 값이 점진적으로 감소하였으며, 이는 모델의 탐지 및 분류 성능이 점차 향상되었음을 의미한다.

본 연구에서 Mask R-CNN을 활용한 간판 탐지 학습 결과는 그림 13과 표 3과 같다. 전체 에폭 동안 학습 손실 값은 초기 약 0.377에서 시작하여 에폭 24에서 0.1 이하로 감소하였으며, 이는 모델의 학습 성능이 지속적으로 개선되고 있음을 나타낸다. 검증 손실 역시 에폭 1에서 약 0.17이었으나, 에폭 24에서는 약 0.02로 감소하여 유사한 경향을 보였다.

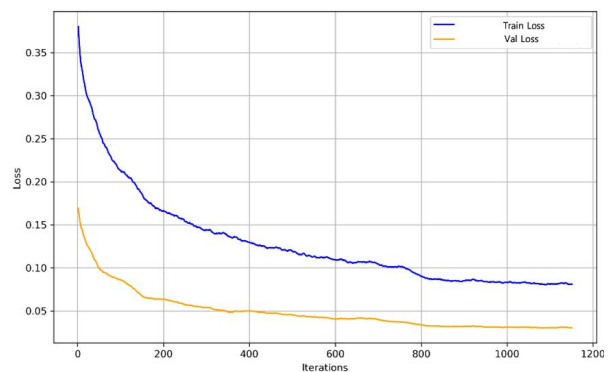


그림 13. 간판 학습 반복 횟수에 따른 검증 및 검증 손실 값 변화

Fig. 13. Training and validation loss trends iterations for sign boards

표 3. Mask R-CNN 모델의 간판 학습시 주요 손실 값
Table 3. Main loss values for the mask R-CNN model when training signs

Iteration	Train loss	Val loss
0	0.377	0.168
200	0.166	0.065
400	0.127	0.050
600	0.112	0.031
800	0.088	0.022
1000	0.086	0.019
1100	0.085	0.018

신분증 학습 정확도와 마찬가지로, 그림 14와 표 4의 간판 학습 정확도의 에폭별 변화의 결과에서도 BBox 및 세그멘테이션정확도는 에폭 증가에 따라 뚜렷한 향상을 보였다. BBox 정는 초기 약 0.553에서 에폭 12에 mAP가 약 0.622에 도달하였고 이후로도 지속적으로 증가하였다. Segmentation 정확도 또한 모델이 학습 데이터를 기반으로 점차 높은 예측 성능을 달성하고 있다.

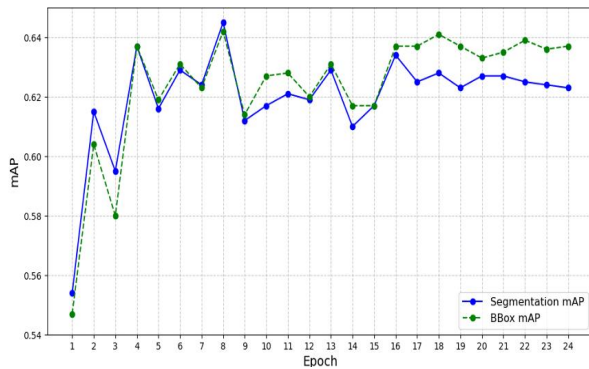


그림 14. 간판 학습 정확도의 에폭별 변화
Fig. 14. Train accuracy over epochs for signboard

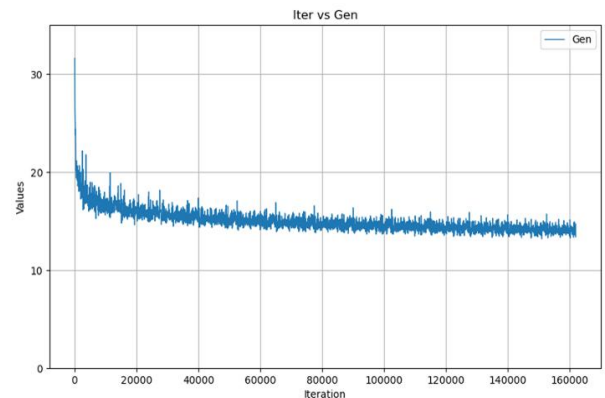
표 4. Mask R-CNN 모델의 간판 학습시 주요 성능
Table 4. Main evaluation score for the mask R-CNN model when training signs

Epoch	Segmentation mAP	BBox mAP
1	0.553	0.545
8	0.645	0.641
24	0.622	0.637

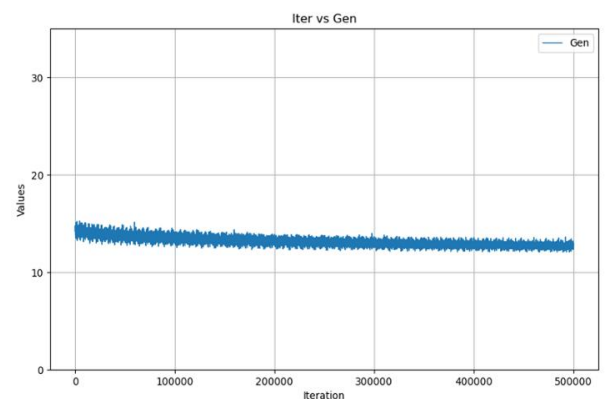
4.3 SRNet 모델 학습 및 평가

SRNet의 한글 텍스트 처리 성능을 향상하기 위

해, 다양한 환경에서 수집된 이미지와 영단어 및 한글 단어를 조합하여 학습용 데이터셋을 구축하고 이를 모델 학습에 활용하였다. 학습 과정에서는 생성 손실 값이 점차 감소하는 경향을 보였으며, 이는 모델의 텍스트 생성 및 편집 성능이 지속적으로 향상되고 있음을 나타낸다. SRNet을 이용한 학습 결과는 그림 15와 같다. 그림 15(a)에 나타난 바와 같이, 초기 10만 개 데이터셋을 학습하는 동안 손실 값은 약 31.6023에서 시작하여 50만 번째 반복(iteration)에서는 약 13.6086으로 14 이하로 감소하였다. 이는 모델이 훈련 데이터를 점차적으로 잘 학습하고 있음을 보여준다. 이후 추가로 10만 개의 데이터를 더 학습하였으며, 41번째 에폭부터 80번째 에폭까지의 학습 결과는 그림 15(b)에서와 같이 점진적인 손실 감소 추세를 유지하였으며 최종적으로 약 12.8268을 기록하였다.



(a) 초기 10만장 반복 학습에서의 손실 변화
(a) Initial 100,000 data loss trends iterations



(b) 10만장 추가 반복 학습에서의 손실 변화
(b) Additional 100,000 data loss trends iterations

그림 15. SRNet 학습에서의 손실 변화
Fig. 15. Loss trends iterations for SRNet

모델의 정답 텍스트와 실제 추론 결과 간의 평균 제곱 오차(MSE, Mean Square Error)는 식 (3)을 통해 평가되었다.

$$MSE = \frac{1}{N} \sum_{i=1}^N (I_o(i) - I_g(i))^2 \quad (3)$$

N 은 사진의 전체 픽셀 수, $I_o(i)$ 는 원본 사진의 i 번째 픽셀 값, $I_g(i)$ 는 생성된 사진의 i 번째 픽셀 값을 의미한다.

그림 16과 표 5에서와 같이, 학습 데이터의 양이 증가함에 따라 MSE는 점진적으로 감소하는 추세를 보인다. 초기 10만 장의 데이터로 40 에폭 동안 학습했을 때 MSE는 약 1,200에서 1,150으로 감소하였으며, 이후 추가로 10만 장의 데이터를 학습한 결과 MSE는 약 1000 수준까지 낮아졌다. 이는 초기 학습 데이터가 한글 위주로 구성된 반면, 추가된 데이터에는 영어 텍스트가 포함되어 있어서 보다 다양한 언어 학습이 가능한 것으로 확인되었다.

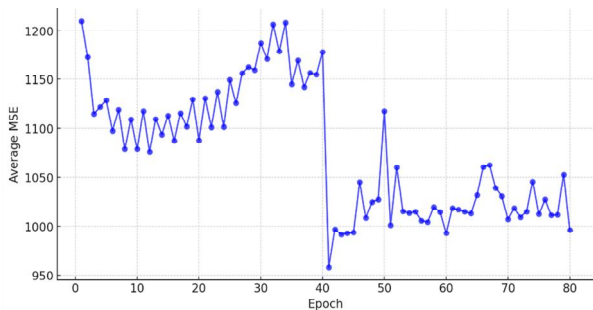


그림 16. SRNet 학습 정확도의 에폭별 변화
Fig. 16. Train accuracy over epochs for SRNet

표 5. Mask R-CNN 모델의 간판 학습시 주요 성능
Table 5. Main evaluation score for the mask R-CNN model when training signs

Epoch	MSE
1	1209.748
40	1178.100
41	957.954
80	996.103

기존 SRNet논문[19]에서 제시하였던 예시 모델의 성능과 본 논문에서 구축한 한글 영어 혼용 데이터

를 활용해 새로 학습한 모델 성능의 차이를 PSNR (Peak Signal-to-Noise Ratio)과 SSIM (Structural Similarity Index Measure) 수치 비교를 통해 확인하였다. 기존 모델에서의 테스트 데이터 500장에 대한 평균 PSNR 값은 16.95 dB, 평균 SSIM 값은 0.5158로 기록되었다. 같은 데이터에 대해서 본 연구에서 학습한 모델의 수치는 평균 PSNR이 18.91 dB, 평균 SSIM은 0.6180으로 더 향상된 수치를 기록하였다. 이러한 결과는 제안된 데이터셋이 다양한 문자 형태와 혼용 양상을 효과적으로 반영했으며, 그로 인해 모델의 일반화 성능이 개선되었음을 나타낸다. 전반적으로 객관적인 수치 지표(PSNR:1.96dB향상, SSIM: 0.1022상승) 모두에서 기존 대비 이미지 품질 복원 측면에서 의미 있는 개선과 구조적 유사성 보존이 우수한 것으로 확인되었다.

4.4 기존 유사 연구와 비교분석

현재 국내외적으로 얼굴 가명화에 대한 연구가 진행되고 있으나 신분증에 대한 이미지 가명화 연구는 보고되지 않아 비교 실험에 제한점이 있다. 기존에 제안된 시스템은 개인정보에 대한 글자인식 및 블러링 처리 기법에 대한 연구가 있으나, 본 연구에서는 개인정보에 대한 가명화 처리 기능을 제공한다. 기존 개인정보시스템과 비교 평가는 표 6과 같다. 첫째, 이름, 전화번호, 주소 등 개인정보 유형에 따라 인식된 텍스트를 명확하게 분류할 수 있으며, 신분증 자동 검출이 적용된 논문과 비교하여 개인정보 인식률은 13.5% 향상되었다. 둘째, 다양한 각도로 왜곡된 이미지에서도 정확한 보정과 탐지가 가능하도록 Convex Hull 기법을 적용하여 비정형적이고 가변적인 데이터의 위치를 정확히 탐지하고 다양한 텍스트 크기에 적응할 수 있다. 셋째, 신분증 개인정보 보호를 위한 기존의 블러링처리 기법을 개선하여 신분증 가명화 기능을 제공한다. 또한 신분증 모델을 간판모델로 확장 가능성을 실험하였다. 전반적으로, 이 시스템은 사형변환(Affine transformation)이 적용된 이미지로부터 94.7%의 성공적인 분류와 자연스러운 이미지의 가명화가 가능함으로 나타났다.

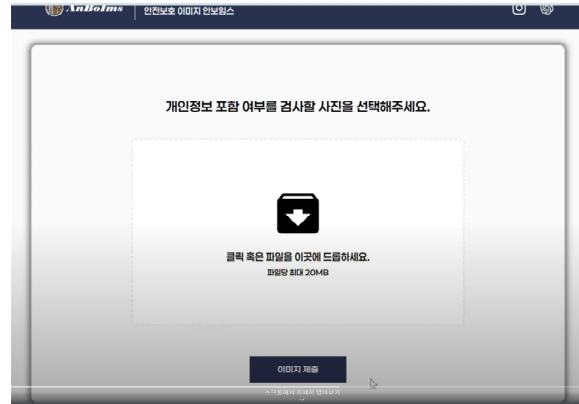
표 6. 유사 모델과 비교 평가 결과

Table 6. Comparison and evaluation results with similar models

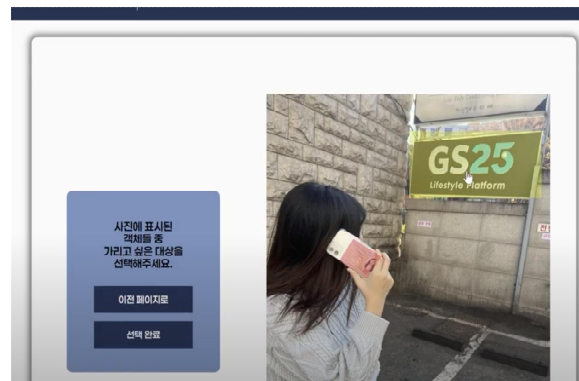
Category	Park[21]	Kim[22]	Proposed model
Types of personal information	ID, addresses, contact numbers, account numbers	Recognizes names, addresses, and contact numbers	Recognizes names, addresses, and contact numbers
Recognition of image distortion	Fixed pattern recognition based on digit position	Recognition based on individual characters in an image Character Region Awareness for Text detection	Recognition based on individual characters in an image Character Region Awareness for Text detection
Applied technology	TesseractOCR + CNN	Transformer-based OCR, CNN, NER, Real-ESRGAN, CRAFT	Mask R-CNN + SRNet
Recognition rate	- For ID numbers 94% - 86.6% recognition of faces	Accurate recognition of personal information 81.2%	Accurate recognition of personal information 94.7%
Specialized area	No specialized model	Specialized model enabled (blurring)	Specialized model enabled (pseudonymization processing)

4.5 모델 구현 결과

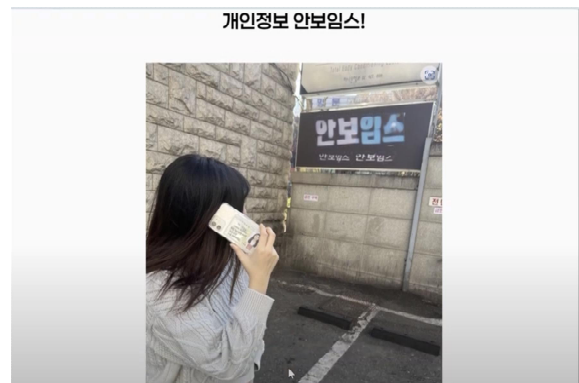
제안한 모델을 실제 현장에 적용한 결과는 그림 17과 같다. 그림 17 (a)는 사용자가 가명화를 원하는 이미지를 업로드하는 초기 단계이며, (b)는 시스템이 개인정보 유출 가능성이 있는 객체를 탐지한 후, 사용자가 가명화할 객체를 선택하는 과정이다. 이후, 그림 17 (c)에서는 선택된 객체에 대해 종류에 따라 적절한 가명화 처리가 수행되고, 결과 이미지가 출력된다.



(a) 사진 업로드
(a) Picture upload



(b) 객체 탐지 결과
(b) Object detection result



(c) 가명화 결과
(c) Pseudonymization result

그림 17. 모델 구현 결과
Fig. 17. Implementation results of the model

그림 18 (a)는 실험용 신분증 객체가 탐지된 실험결과로, 시스템이 이름, 주민등록번호, 주소 등의 개인정보를 포함한 영역을 정확히 인식한 것을 보여준다. 그림 18 (b)에서는 해당 개인정보가 자연스러운 가명화 결과를 확인할 수 있다.



(a) 가명처리 전

(a) Before pseudonym processing



(b) 가명처리 후

(b) After pseudonym processing

그림 18. 신분증 가명처리 결과

Fig. 18. Result of pseudonym processing for Id card

또한, 그림 19 (a)는 상호명과 전화번호가 포함된 간판 이미지이며, (b)는 이를 가명화한 결과로, 기존의 블러링 처리 기법보다 더 시각적으로 자연스러움을 유지하면서 개인정보를 효과적으로 차단하고 있음을 보여준다.



(a) 가명처리 전

(a) Before pseudonym processing



(b) 가명화 결과

(b) Pseudonymization result

그림 19. 간판 가명처리 결과

Fig. 19. Result of pseudonym processing for signboard

V. 결론 및 향후 과제

본 논문에서는 신분증과 간판 등 개인정보가 포

함된 객체를 자동으로 탐지하고, 사진의 배경을 왜곡하지 않고 자연스러운 방식으로 가명화할 수 있는 개인정보 보호 시스템을 제안하고 구현하였다. 제안된 시스템은 Mask R-CNN 기반 객체 탐지 모델과 SRNet 기반 텍스트 변환 기법을 통합하여 다음과 같은 주요 성과를 도출하였다.

첫째, 다양한 환경에서 수집된 신분증 및 간판 이미지 데이터를 바탕으로 Mask R-CNN 모델을 학습시켜, 개인정보 객체를 높은 정확도로 탐지할 수 있다.

둘째, SRNet 기반 텍스트 변환 기법을 통해 비정형 이미지 내 텍스트 정보를 시각적 이질감 없이 자연스럽게 가명 처리할 수 있다.

셋째, 객체 탐지와 가명화 전 과정을 자동화함으로써 데이터의 활용 가치를 훼손하지 않으면서도 효과적인 개인정보 보호를 실현하였다.

실험 결과, 제안 시스템은 SNS와 같은 온라인 플랫폼에서 발생할 수 있는 비의도적 개인정보 유출을 예방하는 데 유의미한 잠재력을 보였으며, 가명화 기반의 신분증 학습 데이터셋 구축에도 활용 가능성을 확인하였다. 이는 향후 관련 연구 및 실무 응용에 있어 실질적인 기여가 될 것으로 기대된다.

본 연구의 시스템은 기존의 단순 익명화 방식과 달리, 데이터의 유용성을 유지하면서도 높은 수준의 개인정보 보호를 제공한다는 점에서 실용적 가치가 크다.

향후 연구에서는 신분증과 간판 외에도 다양한 유형의 개인정보 객체로 탐지 대상을 확장하여 보다 광범위한 데이터 활용 환경에서의 보호 체계를 마련하고자 한다. 또한, 객체 탐지 모델을 이미지-텍스트 매칭 기반의 멀티모달 탐지 모델로 확장하고, 복잡한 배경에서도 자연스러운 가명 처리가 가능하도록 SRNet의 성능을 고도화하는 방안을 모색할 계획이다.

References

- [1] B. Kim, "Privacy-Enhancing Technologies for AI and the Challenges of Legal Frameworks", Korean Journal of Law and Economics, Vol. 20, No. 1, pp. 113-151, Apr. 2023. <https://doi.org/10.46758/kjle.2023.04.20.1.113>.

- [2] Personal Information Protection Commission, <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=9899>. [accessed: Apr. 01, 2025]
- [3] M. Maximov, I. Elezi, and L. Leal-Taixé, "CIAGAN: Conditional Identity Anonymization Generative Adversarial Networks", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 5446-5455, Jun. 2020. <https://doi.org/10.1109/CVPR42600.2020.00549>.
- [4] H. Hukkelås, R. Mester, and F. Lindseth, "DeepPrivacy: A Generative Adversarial Network for Face Anonymization", Advances in Visual Computing, Lecture Notes in Computer Science, Vol. 11844, pp. 565-578, Oct. 2019. https://doi.org/10.1007/978-3-030-33720-9_44.
- [5] M. Rempe, L. Heine, C. Seibold, F. Hörst, and J. Kleesiek, "De-Identification of Medical Imaging Data: A Comprehensive Tool for Ensuring Patient Privacy", arXiv preprint arXiv:2410.12402, Oct. 2024. <https://arxiv.org/abs/2410.12402>.
- [6] V. V. Arlazarov, K. Bulatov, T. Chernov, and V. L. Arlazarov, "MIDV-500: A Dataset for Identity Document Analysis and Recognition on Mobile Devices in Video Stream", Computer Optics, Vol. 43, No. 5, pp. 818-824, Oct. 2019. <https://doi.org/10.18287/2412-6179-2019-43-5-818-824>.
- [7] A. Kotey, T. Gupta, S. Bharuka, A. Singh, N. Ghugare, and L. Samanthapuri, "Preserving Privacy in Multimedia: Text-Aware Sensitive Information Masking for Visual Data", International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Vol. 10, No. 1, Jan. 2024. <https://doi.org/10.32628/CSEIT2410117>.
- [8] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN", Proc. of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp. 2961-2969, Oct. 2017. <https://doi.org/10.1109/ICCV.2017.322>.
- [9] E. D. Cherpanath, P. R. F. Nasreen, K. Pradeep, M. Menon, and V. S. Jayanthi, "Food Image Recognition and Calorie Prediction Using Faster R-CNN and Mask R-CNN", Proc. of the 9th International Conference on Smart Computing and Communications (ICSCC), Kochi, Kerala, India, pp. 83-89, Aug. 2023. <https://doi.org/10.1109/ICSCC59169.2023.10335053>.
- [10] C. Wang and C. Zhong, "Adaptive Feature Pyramid Networks for Object Detection", IEEE Access, Vol. 9, pp. 107024-107032, Jul. 2021. <https://doi.org/10.1109/ACCESS.2021.3100369>.
- [11] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee, "Character Region Awareness for Text Detection", Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, pp. 9365-9374, Jun. 2019. <https://doi.org/10.1109/CVPR.2019.00959>.
- [12] H. Kim, S. Kim, M. Kim, S. Kim, Y. Cho, E. Choi, and G. Yoo, "Development of a Named Entity Recognition-based Text Masking System for Preventing Personal Information Exposure in Shipping Labels", The Journal of Korean Institute of Information Technology, Vol. 22, No. 11, pp. 181-196, Nov. 2024. <https://doi.org/10.14801/jkiit.2024.22.11.181>.
- [13] H. Aljzaere, A. C. Nath, and W. Hardt, "German License Plate Recognition System Using the YoloV8 Model and EasyOCR", Proc. of the 2024 IEEE CPMT Symposium Japan, Kyoto, Japan, pp. 176-179, Mar. 2024. <https://doi.org/10.1109/ICSJ62869.2024.10804754>.
- [14] D. R. Vedhaviyassh, R. Sudhan, G. Saranya, M. Safa, and D. Arun, "Comparative Analysis of EasyOCR and TesseractOCR for Automatic License Plate Recognition Using Deep Learning Algorithm", Proc. of the 6th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, pp. 966-971, Jun. 2022. <https://doi.org/10.1109/ICECA55336.2022.10009215>.

- [15] C. Kim, R. Park, and E. Hong, "Breast Mass Classification Using eLFA Algorithm Based on CRNN Deep Learning Model", IEEE Access, Vol. 8, pp. 197312-197323, Oct. 2020. <https://doi.org/10.1109/ACCESS.2020.3034914>.
- [16] J. Zhu, T. Park, P. Isola, and A. Efros, "CycleGAN: Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks", Proc. of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp. 2223-2232, Oct. 2017.
- [17] P. Isola, J. Zhu, T. Zhou, and A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks", Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, Hawaii, pp. 1125-1134, Jul. 2017.
- [18] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. Huang, "Generative Image Inpainting with Contextual Attention", Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, pp. 5505-5514, Jun. 2018.
- [19] L. Wu, C. Zhang, J. Liu, J. Han, J. Liu, E. Ding, and X. Bai, "Editing Text in the Wild", arXiv preprint arXiv:1908.03047, Aug. 2019. <https://doi.org/10.48550/arXiv.1908.03047>.
- [20] P. Roy, S. Bhattacharya, S. Ghosh, and U. Pal, "STEFANN: Scene Text Editor Using Font Adaptive Neural Network", Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, pp. 13228-13237, Jun. 2020.
- [21] S. C. Park, "Design and Implementation of Personal Information Identification and Masking System Based on Image Recognition", The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 17, No. 5, pp. 1-8, Oct. 2017. <https://doi.org/10.7236/JIIBC.2017.17.5.1>.
- [22] H. E. Kim, S. Y. Kim, M. S. Kim, S. H. Kim,

Y. J. Cho, E. S. Choi, and G. S. Yoo, "Development of a Named Entity Recognition-based Text Masking System for Preventing Personal Information Exposure in Shipping Labels", The Journal of Korean Institute of Information Technology, Vol. 22, No. 11, pp. 181-196, Nov. 2024. <https://doi.org/10.14801/jkiit.2024.22.11.181>.

저자소개

한 영 옥 (Younguk Han)



2024년 12월 : 고려대학교
지능정보 SW아카데미 5기
수료(640H)
2025년 2월 : 숭실대학교
경영학부/AI융합학부(학사)
2025년 2월 ~ 현재 : T3Q(주) 주임
관심분야 : 컴퓨터비전, 딥러닝,
LLM

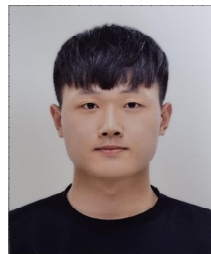
우 혜 진 (Hyejin Woo)



2024년 12월 : 고려대학교
지능정보 SW 아카데미 5기
수료(640H)
2021년 3월 ~ 현재 : 고려대학교
세종캠퍼스
인공지능사이버보안학과
학사과정

관심분야 : 컴퓨터 비전, 딥러닝, LLM, Multi-Modal

우 지 호 (Jiho Woo)



2024년 12월 : 고려대학교
지능정보 SW아카데미 5기
수료(640H)
2019년 3월 ~ 2025년 2월 :
고려대학교 체육교육과(학사)
2025년 3월 ~ 현재 : 고려대학교
통계학과 석사과정

관심분야 : 딥러닝, 컴퓨터비전, 데이터사이언스,
데이터분석

김 규 리 (Kyuri Kim)



2024년 12월 : 고려대학교
지능정보 SW아카데미 5기
수료(640H)
2022년 3월 ~ 현재 :
중앙대학교 소프트웨어학부
학사과정
관심분야 : 딥러닝, 컴퓨터비전,
LLM, 클라우드 컴퓨팅

이 수 정 (Sujung Lee)



2024년 12월 : 고려대학교
지능정보 SW아카데미 5기
수료(640H)
2021년 3월 ~ 현재 : 고려대학교
세종캠퍼스
컴퓨터융합소프트웨어학과
학사과정

관심분야 : 컴퓨터비전, 딥러닝, 자율주행

유 길 상 (Gilsang Yoo)



2010년 2월 : 중앙대학교
영상공학과(박사)
2010년 3월 ~ 현재 :
(사)한국컴퓨터게임학회 부회장
2011년 3월 ~ 현재 : 고려대학교
정보대학 정보창의교육연구소/
지능정보 SW아카데미 교수

2023년 3월 ~ 현재 : (사)한국미디어 아트산업협회
수석부회장

관심분야 : 데이터사이언스, 데이터 시각화, 빅데이터
분석, 3D영상 콘텐츠, 머신러닝, 딥러닝, 컴퓨터교육