

모바일 환경에서의 이미지 내 텍스트 마스킹 기반 개인정보 보호 시스템

오희주, 김상대

순천향대학교 의료IT공학과

ohj5018@sch.ac.kr, sdkim.mie@sch.ac.kr

A privacy Protection System Based on Image Text Masking in Mobile Environments

Huiju Oh, Sangdae Kim

Dept of Medical IT Engineering, Soonchunhyang University

요약

SNS에 민감한 정보가 포함된 이미지를 사용자가 인지하지 못한 채 게시하여 정보가 유출되는 사례가 발생하고 있다. 본 연구에서는 광학문자인식(OCR) 기술을 활용하여 이미지 내 텍스트 정보를 자동으로 감지하고, 사용자가 민감하다고 판단한 영역만 선택하여 마스킹하는 Android 기반 개인정보 보호 시스템을 제안한다. 감지된 텍스트는 시각적으로 표시되며, 선택한 영역에만 마스킹이 적용되어 불필요한 정보 손실 없이 효율적인 보호가 가능하다. 이를 통해 사용자가 놓치기 쉬운 민감 정보까지 효과적으로 차단하여 정보 유출을 예방할 수 있다.

I. 서론

현대 사회에서 SNS, 블로그 등을 통한 이미지 공유는 개인의 일상을 기록하고 타인과 소통하는 주요 수단으로 활용되고 있다. 그러나 이러한 이미지 공유 과정에서 위치정보, 이름, 주민등록번호, 주소 등과 같은 민감한 개인정보가 텍스트 형태로 포함되는 경우가 발생한다.[1] 이는 단순한 사생활 침해를 넘어 범죄에 악용될 가능성도 존재하여 심각한 보안 문제로 이어질 수 있다.[2]

특히 이미지 내에 텍스트 기반 개인정보는 얼굴이나 신체 이미지보다도 개인을 더 명확하게 식별할 수 있음에도 불구하고, 마스킹하는 작업에서 사용자가 인지하지 못한 채 이미지를 게시하여 마스킹이 누락되는 경우가 발생하기도 한다.[3] 현재 대부분의 사용자는 이미지 편집 도구를 활용하여 민감한 정보를 수작업으로 마스킹하고 있으나, 이는 번거로운 뿐 아니라 정보량이 많은 경우 어떤 항목을 가려야 할지 판단하기 어렵다.

실제로 보안뉴스에 따르면, 한 보험설계사가 네이버 밴드에 고객의 신분증과 카드 정보가 포함된 이미지를 게시한 사건이 있었고, 이후 해당 정보가 악용되어 금융 사기로 이어진 사례가 보고된 바 있다.[2] 이처럼 텍스트 기반 개인정보는 매우 민감함에도 불구하고, 기존 시스템은 사용자의 의견과 상관없이 일괄적으로 마스킹을 적용해, 불필요한 정보까지 가려지는 문제가 발생하고, 이로 인해 사용자의 의도를 반영하기 어렵다는 한계가 있다.[4]

본 논문에서는 이러한 문제를 해결하고자 이미지 내 텍스트 정보를 자동으로 감지하고, 이를 사용자에게 시각적으로 제공함으로써 사용자가 민감하다고 판단하는 텍스트만을 직접 선택하여 마스킹할 수 있는 Android 기반 개인정보 보호 시스템을 제안한다. 이를 통해 사용자는 사전에 인식하지 못했던 정보까지 확인하고 보호할 수 있으며, 마스킹 과정의 효율성을 높여 개인정보 유출을 사전에 방지할 수 있을 것으로 기대된다.

II. 본론

본 장에서는 제안하는 시스템의 전체적인 구조와 기술적 구현 방식에 대해 설명한다.

II-1. 기존 연구

기존의 이미지 기반 개인정보 보호 도구들은 이미지 내 텍스트를 모두 일괄적으로 모자이크 처리를 적용하는 방식으로 구현되어 있다.[4] 이러한 방식은 사용자의 의도를 반영하기 어려우며, 비민감한 정보까지도 불필요하게 마스킹 되는 한계가 있다. 또한 수작업으로 민감한 정보를 식별하여 마스킹하는 과정은 번거롭고, 정확하고 빠른 작업이 어렵다.

본 연구는 이러한 문제를 해결하고자, 이미지 내 텍스트 정보를 자동으로 인식하는 광학문자인식(OCR) 기술, 사용자의 선택에 기반한 마스킹 처리 방식, 그리고 실시간 시각적 피드백 기능을 결합한 Android 기반 개인정보 보호 시스템을 구현하는 것을 목표로 한다.

II-2. 시스템 구조

본 시스템은 Android 기반의 클라이언트 애플리케이션과 Flask기반의 Python 서버로 구성된다. 사용자는 앱을 통해 이미지를 선택하거나 촬영하며, 해당 이미지는 서버로 전송된다. 서버는 수신된 이미지를 Tesseract OCR을 통해 분석하고, 이미지 내 텍스트의 문자열과 위치 좌표를 추출하여 JSON 형식으로 클라이언트에 반환한다.

클라이언트는 수신된 정보를 바탕으로 텍스트 영역을 시각적으로 표시하고, 사용자는 마스킹이 필요한 영역을 선택한다. 선택된 좌표 정보는 서버로 다시 전송되며, 서버는 OpenCV를 이용해 해당 영역에 Gaussian Blur를 적용한 마스킹 이미지를 생성하여 클라이언트에 반환한다.

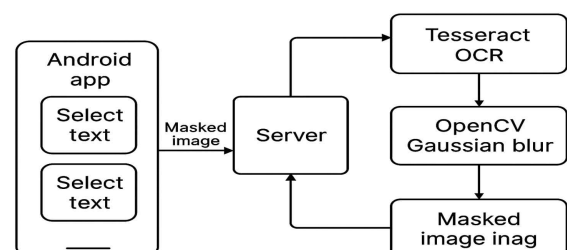


그림 1. 시스템 전체 구조도

III. 시스템 동작과정

본 시스템은 클라이언트와 서버 간의 통신을 통해 이미지 내 텍스트 감지 및 선택적 마스킹 기능을 수행한다. 본 절에서는 사용자가 갤러리에서 이미지를 선택했을 때를 기준으로 시스템의 동작 과정을 설명한다.

사용자가 갤러리에서 이미지를 선택하면, 클라이언트는 해당 이미지를 Retrofit2 라이브러리를 통해 HTTP POST 방식으로 서버에 전송한다. 서버는 수신된 이미지를 OpenCV로 전처리한 뒤, Tesseract OCR을 이용하여 이미지 내 텍스트를 감지하고, 각 텍스트의 문자열과 위치 정보를 추출한다. 감지 결과는 JSON 형식으로 구성되어 클라이언트에 전달된다.

클라이언트는 수신된 정보를 바탕으로 이미지상에 감지된 텍스트 영역을 사각형 박스로 시각화하며, 사용자는 이 중 민감하다고 판단되는 영역을 선택하여 마스킹 대상으로 지정할 수 있다. 선택된 영역의 좌표 정보는 다시 서버로 전송되며, 서버는 해당 영역에 Gaussian Blur를 적용한 마스킹된 이미지를 생성하고 이를 클라이언트에 반환한다.

이 과정을 통해 사용자는 자신이 미처 인지하지 못한 민감한 정보까지 효과적으로 식별할 수 있으며, 직관적인 UI를 통해 간편하게 마스킹 작업을 수행할 수 있다.

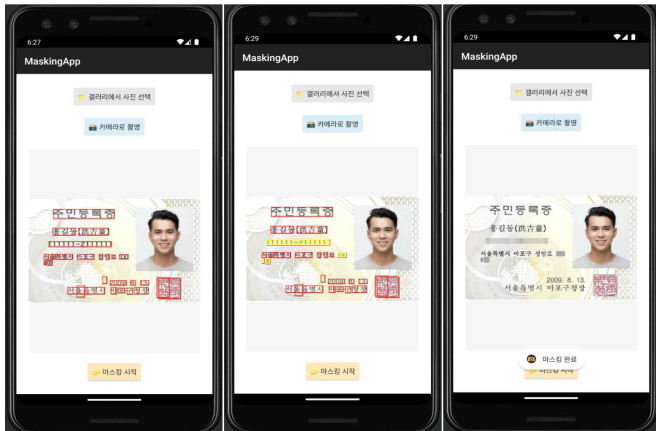


그림 2. 텍스트 자동 감지 결과 그림 3. 민감 정보 선택 화면 그림 4. 마스킹 처리 완료 화면

IV. 결론

본 논문에서는 이미지 내 텍스트 정보를 자동으로 식별하고, 사용자가 보호가 필요하다고 판단한 항목만을 선택적으로 마스킹할 수 있는 Android 기반의 개인정보 보호 시스템을 제안하였다. 제안된 시스템은 기존의 일괄 마스킹 방식과 달리, 사용자의 의견을 반영하여 필요한 정보만을 가릴 수 있도록 하여 정보 전달력을 유지하면서도 개인정보 노출을 효과적으로 방지할 수 있도록 설계되었다.

Tesseract OCR을 활용하여 이미지 내 텍스트를 자동으로 인식하고, 감지된 정보를 시각적으로 표현함으로써 사용자가 직관적으로 마스킹 대상을 인식할 수 있도록 하였다. 이를 통해 민감한 정보의 누락 가능성을 줄이고, 사용자의 인지 기반의 능동적인 개인정보 보호를 실현할 수 있다.

향후 연구에서는 Gaussian Blur 외에도 모자이크 처리, 텍스트 제거 등 다양한 마스킹 기법을 추가하여 사용자의 선택 폭을 넓히고자 한다. 또한, 민감한 정보로 판단되는 항목을 자동으로 추천하는 기능을 추가함으로써 사용자의 판단 부담을 줄이고 보안성을 강화할 수 있을 것으로 기대된다.

본 시스템은 선택 기반의 마스킹 기능을 통해 이미지 공유가 활발한 모바일 환경에서 발생할 수 있는 개인정보 유출 문제를 효과적으로 예방할

수 있으며, 다양한 이미지 기반 플랫폼과 보안 시스템에 적용될 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 2025년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구 결과로 수행되었음(2021-0-01399)

참 고 문 헌

- [1] "인터넷에 올린 사진 보고 주소 알아내면 어떡하죠" 불안한 여성들, 아시아경제, 임주형 기자, <https://www.asiae.co.kr/article/2021050612335659939>
- [2] SNS상 개인정보 이용해 피해자 통장에서 1억원 탈취한 범인 잡혔다, 보안뉴스, 김영명 기자, <https://www.boannews.com/media/view.asp?idx=111126>
- [3] [PIS FAIR 2022] 엘세븐시큐리티, 이미지 속 개인정보 유출사례, 보안뉴스, 김영명 기자, <https://www.boannews.com/media/view.asp?idx=107247>
- [4] 오희주, 김상대, (2024-06-19). 개인정보 보호를 위한 위치정보 마스킹 시스템. 한국통신학회 학술대회논문집, 제주.