

스마트시티를 위한 시각데이터의 활용

김선호 (University of Southern California)

목 차	1. 서 론
	2. 시각데이터 통합 플랫폼
	3. 시각데이터 이용 사례
	4. 결 론

1. 서 론

세계적인 도시화의 진행으로 2025년에는 전세계 35개 도시가 천만 명 이상의 인구를 가질 것으로 예상되고 있다. 도시 인구의 증가로 교통혼잡과 같은 도시 문제가 심화될 뿐만 아니라 에너지 생산 증가에 따른 환경 문제도 악화되고 있다. 이미 도시들은 모든 에너지의 3분의 2 이상을 소비하고 다른 자원들의 대부분도 소비하고 있다. 이에 따라 도시화의 문제를 해결하면서 지속 가능한 개발과 동시에 도시 삶의 질을 향상시키기 위해 정보통신(ICT), 사물인터넷(IoT)과 같은 기술을 이용한 스마트시티 사업이 전 세계적으로 관심을 받으며 투자를 받고 있다. 스마트시티는 다양한 방법으로 데이터를 수집하고 분석하여 자산과 자원을 효율적으로 관리하고 활용하는 응용을 총칭하는데 이는 교통망, 에너지, 상하수도, 청소, 환경, 헬스케어 등과 같은 도시 운영 및 서비스 효율성의 최적화는 물론 시민과의 정

보공유나 협력 등도 포함하는 광범위한 개념이다. 따라서, 스마트시티는 선진국이나 개발도상국 할 것 없이 전 세계 도시들이 공통적으로 추진하고 있는 사업이다. 기술을 이용한 새로운 인프라의 구축과 같은 글로벌 스마트시티 시장은 2020년까지 약 1조6000억 달러(약 1810조원)에 이를 것으로 전망되며 점점 관련 시장이 확대되고 있다. 또한 최근에는 인프라 구축과 같은 물리적 측면을 넘어 환경, 생활, 경제, 거버넌스와 같은 비물리적인 측면까지 스마트시티의 개념이 확대되고 있어 미래 산업과 생활의 핵심으로 자리 잡고 있다.

스마트시티는 과학기술 이외에도 지역사회 및 구성원의 참여가 필요하고, 그들의 삶까지 결합된 복잡한 생태계를 요구한다. 기술적으론 도시 구석구석을 자동으로 실시간 모니터링할 수 있는 센서와 통신 기술, 수집된 대용량의 데이터를 처리할 수 있는 빅데이터와 데이터 분석기술이 필요하지만 이런 일련의 행위에 동참하고 협력

하는 집단지성과 같은 도시 구성원의 참여와 유기적인 협력이 스마트시티 성공의 필수요소다. 특히 미래 도시생활의 커다란 변화를, 정부가 주도하는 중앙화된 사업을 수동적으로 따라가는 것만이 아니라 도시 구성원들이 능동적으로 스마트시티 사업에 참여해 같이 만들어가는 것이 중요하다. 이는 막대한 투자가 필요한 미래 도시의 구축에 있어 최소의 자원으로 최대의 결과를 만들어 내기 위함이다. 스마트시티 추진은 또한 많은 기회를 제공하기도 한다. 신기술을 앞세운 기업들 뿐만 아니라 기존 기술의 새로운 응용 기회도 다양한 분야에서 발생하기 때문에 관심을 갖고 주시해야 한다.

스마트시티의 기반이 되는 대규모 모니터링에 있어 인공지능의 발전은 거리에서 발생하는 수많은 일들을 카메라를(CCTV, 블랙박스, 스마트폰 등) 이용해 이미지나 비디오와 같은 시각데이터로 수집 후 이미지 인식 기능을 활용해 더 많이 모니터링을 자동으로 할 수 있음을 의미한다. 이미 이미지 인식 기술은 전례 없는 속도로 엄청난 양의 이미지 데이터를 처리할 수 있고 단순한 패턴 인식을 넘어 기계학습을 통해 빠르게 진보하고 있다. 또한 예전엔 시도되지 않았던 다양한 물체인식을 제공하게 되었고 도시는 이런 이미지 인식을 사용하여 사람과 사물의 움직임을 더 잘 이해할 수 있게 되었다. 풍부한 데이터를 자유로이 사용할 수 있게 된 도시 계획자와 시 공무원들은 최적의 안전과 효율성을 위해 도시의 물리적 인프라를 개선할 방법에 대해 보다 효과적인 결정을 내릴 수 있게 되었다. 본 논문은 스마트시티 응용에 있어 도시 커뮤니티 구성원들이 시각데이터를 보다 효과적으로 활용할 수 있는 방법을 제안하고 미국 로스앤젤레스시의 실제 데이터 세트를 이용한 사례를 설명한다.

2. 시각데이터 통합 플랫폼

이미지 및 비디오와 같은 시각데이터는 이미 널리 사용되고 있고, 그 자체가 새로운 커뮤니케이션 소스를 형성하며, 또한 스마트시티 응용을 위한 정보 및 지식을 제공하고 있다(예: 거리 청결도 조사[1], 교통 흐름 분석[2], 재난 상황 인식[3], 도로면 손상 감지[4], 물체 인식[5] 및 시각 정보 생성[6,7]). 특히 모바일 카메라(스마트 폰, 드론, 차량 블랙 박스 등)는 다양한 목적으로 도시 구석구석을 기록하고 있고, 이러한 시각데이터 자체가 도시 생활에 대해 풍부한 정보를 담고 있는 중요한 자산이 되고 있다. 더 나아가, 급속도로 진보하는 기계학습이나 인공지능같은 기술들이 시각데이터를 자동으로 분석하고 숨겨진 새로운 유용한 정보와 지식을 추출해내기에 그 중요성이 더욱 부각되고 있다.

인프라로서의 시각데이터는 스마트시티에서 큰 잠재력을 약속하며 간단한 낙서신고 응용(예: MyLA311 App[8])에서부터 지능형 테러감지 시스템과 같은 정교한 작업에 이르기까지 광범위한 범위에서 도시의 기능을 향상시키에 많은 기관에서 다양한 시각데이터 애플리케이션을 개발하고 있다. 최근 딥러닝과 같은 기계학습 기술의 급속한 발전은 공공데이터를 누구나 쉽게 민주적 방식[9]으로 활용할 수 있는 좋은 기회를 제공한다. 즉, 기계학습과 인공지능은 새로운 가치 창출에 엄청난 잠재력을 가지고 있고, 훌륭한 것을 창조할 수 있는 지능을 가진 개인이나 단체들은 어디에서나 나타날 수 있기에, 이런 사람들이 그들의 잠재력을 최대한 발휘할 수 있도록 필요한 데이터, 지식 및 분석 도구에 쉽게 접근해 사용할 수 있게 만들 필요가 있다.

시각데이터를 인프라로 생각할 때 시각데이터

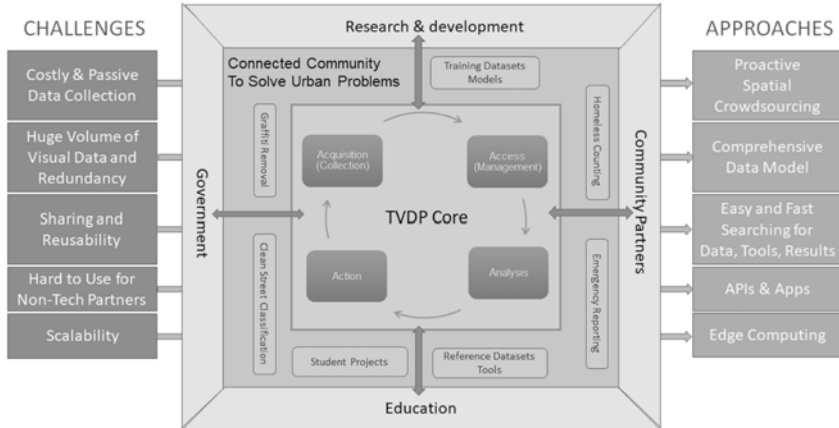
의 사용을 민주화하는 것은 합리적인 전략이 된다. 그러나 시각데이터의 관리에 있어, 특히 자동 수집, 관리 및 분석과 결부해서 여전히 해결해야 할 많은 문제가 있다. 그 중 한 가지 중요한 관찰은 지금까지는 대부분의 응용 프로그램과 접근 방식이 매우 개별적으로 세분화되어와서 유사한 관심사를 가진 커뮤니티 구성원들이 함께 사용하고 공유할 수 있는 체계적이고 시너지 효과가 있는 방법이나 인프라없이 각각의 문제에 대한 개별 솔루션을 제공해왔다는 점이다. 예를 들어, 더 큰 전체에 통합되지 않더라도 각각의 조직이나 기업은 어느 정도 개별 기계학습 애플리케이션의 이점을 얻을 수 있었다. 그러나 이러한 응용 프로그램이 통합 플랫폼 상에 있을 때 이점은 훨씬 더 커지게 된다[10]. 시각데이터의 수집 및 분석(예: 딥 러닝) 비용이 매우 높다는 점을 고려할 때, 특히 대도시 규모의 응용 프로그램에서, 사용자가 수집된 데이터와 분석 도구를 플랫폼에서 공유하여 상호 이익을 위해 공동으로 협력하는 것이 중요하다. 통합 플랫폼의 이점은 참여자가 이를 사용하여 기계 학습에 필요한 데이터에 대한 단일 액세스 지점을 만들 수 있다는 것이다. 일반적으로 데이터 세트가 크고 풍부할수록 학습의 결과가 더 좋아진다. 또한, 잘 설계된 플랫폼은 데이터의 가용성 및 재사용 가능성, 데이터 관리 및 분석 방법의 유용성을 최대화하고, 또 쉽고 효과적인 작업 환경에서 중복 데이터, 분석 작업 및 운영 비용을 최소화할 수 있다.

따라서, USC IMSC 연구소에서는 도시 시각 데이터를 수집, 관리, 분석할 수 있는 통합 데이터 플랫폼을 제안했다[14]. 이 플랫폼을 통해 사용자는 시각데이터 수집, 액세스, 도구, 분석 방법 및 결과를 효율적으로 공유하여 개별 운영을 향상시킬 뿐만 아니라 상호 이익을 얻을 수 있다. 특히 제안된 플랫폼은 이미지가 기록된 곳의 위

치 정보가 대부분의 스마트시티 애플리케이션에서 중요하며 협업자들 간의 데이터 관리 및 공유에 있어 근본적인 연결점을 제공하기 때문에 지리적 태그(spatial tag)가 지정된 시각데이터에 중점을 둔다. 또한, 우리의 연구는 다가오는 기계학습 및 인공지능 애플리케이션 시대의 사용자를 대비하기 위해 고효율 시각데이터 공유 플랫폼 뿐만 아니라 이미지 기반 기계학습 플랫폼을 목표로 한다. 통합 플랫폼은 도시 생활의 기능과 질을 향상시키기 위해 참여 사용자들 사이에서 다양한 이미지 기반 스마트시티 애플리케이션을 파일럿, 테스트 및 실제 적용하는 데 사용될 것이다. 데이터 및 수집 방법의 공유 뿐만 아니라 분석 방법 및 결과도 제공하여 하나의 응용 프로그램에서 사용된 새로운 데이터, 방법 및 추출된 지식을 다른 응용 프로그램으로 효과적으로 변환하여 사용하는 것을 가능하게 하여 궁극적으로 시각데이터 및 분석을 도시 인프라로 만들 수 있다. 우리가 제안하는 통합 플랫폼의 목표는 시각데이터와 그 분석을 통해 광범위하게 사용 가능한 가치 창출을 촉진시켜 스마트시티를 위한 사회 경제적 문제 해결을 커뮤니티 파트너들 간에 보다 분산되고 협력적으로 만드는 것이다.

이 목표를 달성하기 위해 우리가 제안한 시각데이터 통합 플랫폼은 네 가지 핵심 서비스를 제공한다(그림 1).

- 1) 수집: 시각데이터 업로드 및 공간 크라우드 소싱(spatial crowdsourcing)을 포함한 전체적인 데이터 수집 방법,
- 2) 액세스: 시각데이터의 보관, 인덱싱 및 검색을 포함한 데이터 관리 방법,
- 3) 분석: 기계학습을 포함한 시각데이터 분석 방법 및 도구,
- 4) 액션: 분석 결과를 실생활에 적용. 또한, 학습에 의해 개발된 분석 모델을 에지 컴퓨팅



(그림 1) 시각데이터 통합 플랫폼

구성 요소 기능에 따라 다양한 예지 디바이스에 적용하고 예지에서 생성된 결과를 재사용하여 분석 모델을 향상.

이 네 가지 핵심 서비스는 대부분의 데이터 중심 애플리케이션에 적용할 수 있는 데이터 라이프라인 사이클을 형성한다. 사용자는 핵심 서비스의 전부 또는 일부를 이용하여 자신의 목적을 또는 협업을 위해 플랫폼을 활용할 수 있다. 이미지를 빠르게 검색하는 것(예: 액세스만 사용)만큼 간단하거나, 스마트 카메라 센서가 있는 IoT 기반 도시 전체 모니터링 애플리케이션(즉, 전체 사이클을 지속적으로 사용)만큼 복잡할 수도 있다. 특히, 이러한 핵심 서비스는 다양한 참여자들이 스마트시티 애플리케이션을 개별적으로 구현하거나 참여자들 사이에서 공동으로 구현하는 데 사용할 수도 있다. 예를 들면, 다음과 같은 참여자들이 있을 수 있다: 1) 정부: 관심있는 문제에 대한 개방형 데이터 세트를 제공, 2) 전문 연구자 및 개발자: 알고리즘 및 기술 솔루션을 제공, 3) 커뮤니티 파트너: 기술 솔루션을 운영하거나 새로운 데이터를 클라우드 소싱, 4) 학계: 분석 모듈을 구축하거나 확장하기 위해 개방형

데이터 세트를 사용.

일반적으로 시각데이터 플랫폼에는 다음과 같은 이슈들이 있다. 1) 수동 데이터 수집은 비용이 높으며 특정 학습을 달성하기에 충분한 품질의 데이터 세트를 제공하기 힘들다. 2) 시각데이터는 크기가 크고 중복이 많아서 효율적으로 관리하고 검색하기가 매우 어렵다. 3) 데이터 수집 방법, 저장, 액세스 방법에 있어 공동 작업자 간에 재사용 가능한 리소스로 처리를 공유하는 체계적인 방법이 별로 없다. 4) 기술에 익숙하지 않은 파트너가 수집된 데이터와 방법을 자신의 목적으로 사용하기가 쉽지 않으며, 5) 대규모 도시 응용 프로그램에서 확장 가능한 시스템이 필요하다. 이러한 문제를 극복하기 위해 우리가 제안한 플랫폼은 다음의 접근 방법을 제공한다(그림 1).

- 공간 크라우드 소싱(spatial crowdsourcing): 참가자가 모바일 카메라를(예: 스마트폰) 이용하여 특정 위치에서 특정 유형의 시각데이터에 대한 데이터 수집 캠페인을 만들 수 있도록 하여 사전에 양질의 데이터를 수집.
- 수집된 데이터의 포괄적인 모델링: 대상 분석 및 다양한 유형의 향후 분석을 위해 수집된 데이터의 유용성을 지원하는 공간적

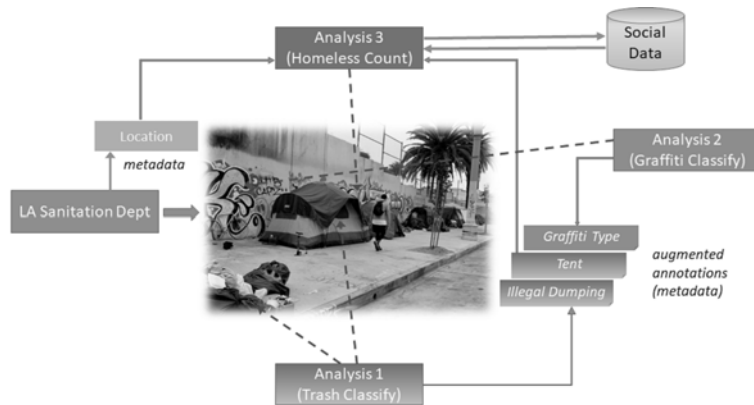
(spatial), 의미적(semantic), 시각적(visual) 기능을 포함한 다양한 유형의 메타데이터를 활용.

- 효과적인 데이터 관리 방법: 포괄적인 데이터 모델을 지원하기 위해 관리 시스템에서 시각데이터를 저장, 액세스 및 검색. (특히, 위치를 이용한 데이터 통합)
- API 및 샘플 앱: 플랫폼의 간편한 사용을 지원하고 경제적이고 민주적인 방식으로 애플리케이션 또는 프로토타입을 신속하게 개발하도록 지원.
- 에지 컴퓨팅: 대규모 스마트시티 플랫폼에서 필수적인 확장성을 위해 컴퓨팅 집약적인 대용량 시각데이터 처리 및 기계학습을 에지 디바이스로 분산.







3. 시각데이터 이용 사례

본 연구에서 사용된 데이터는 스마트시티 애플리케이션을 위해 위치정보가 태그된 시각데이터이다. 따라서 기본 대상 데이터 세트는 위치 메타데이터가 있는 거리 이미지와 비디오로 구성된다. 예시 시나리오는 다음과 같다 (그림 2).

- 1) 정부기관인 LASAN(Los Angeles Sanitation Department)은 스마트폰으로부터 쓰레기가 놓인 거리 장면의 시각데이터를 수집하고 업로드한다(개별 데이터 수집).
- 2) 수집된 이미지와 비디오는 공간적 시간적으로 구성되어 거리 청소를 위한 특정 물체(즉, 불법 투기, 버려진 가구, 노숙자 야영 등 거리에 있는 물체를 인식하기 위해 구축된 분류 모델, 그림 3 참조)를 자동으로 탐지하도록 분석된다(데이터 액세스 및 분석).
- 3) 분류된 결과 및 위치는 신속한 거리 청소(조치)를 위해 LASAN에 보고되고 통합 플랫폼 데이터베이스에 증강된 지식으로 다시 저장된다(액션).
- 4) 분류 결과 중 하나인 노숙자 야영지의 위치와 이미지를 로스앤젤레스시의 다른 정부기관인 노숙자 코디네이터가 독립적으로 액세스하여 노숙자 텐트의 위치를 추적하여(공유 분석 결과) 로스앤젤레스시에서의 텐트 밀집도, 노숙자의 시공간적 이동(사회적 영향)등 사회 문제 해결을 위한 별도의 연구 수행이 가능하게 한다.
- 5) 한 부서에서 수집한 데이터와 분석 결과들



(그림 2) 시각데이터 이용 사례

Image Label	Description	Examples	Image Label	Description	Examples
Bulky Item – Few	1 to 3 items (e.g., coach, desk, mattress, and tire) are thrown on the street.		Encampment	A tent for people who live in streets.	
Bulky Item – Many	More than 3 items are thrown on the street.		Overgrown Vegetation	There is extra vegetation on the streets.	
Illegal Dumping	There is an area which is full of waste which needs special equipment to remove.		Clean	The street is clean 😊	

(그림 3) 거리 청소를 위한 특정 물체 분류

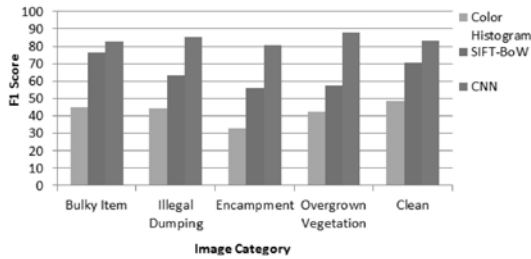
이 다른 부서에서 다른 목적으로 용이하게 이전되어 분석할 수 있어(예: 낙서 감지, 거리 청결 및 안전 수준 간의 상관관계 연구) 학습 시간과 노력을 절약하게 한다.

3.1 거리 청소

대도시의 거리 청결은 도시 환경과 건강에 큰 영향을 미치므로 도시는 거리를 깨끗하게 만들기 위해 많은 노력을 기울이고 있다. 최근 기술의 발전으로 거리 청결도를 대규모로 모니터링하기 위한 스마트 시스템을 개발하는 것이 가능해졌다. USC IMSC 연구소에서는 로스엔젤레스시와 협력하여 거리 이미지를 학습하여 자동으로 거리의 청결도 수준을 식별하고 쓰레기 물체를 식별하는 연구를 실행했다[1]. 시민이 거리에 놓여진 쓰레기를 사진으로 찍어 신고하면 그것을 자동으로 인식하고 분류하여 쓰레기의 위치와 종류를 이용해 적절한 조치를 신속하게 취할 수 있게 하는 응용이다¹⁾. 이를 위해 LASAN에서 수집한 큰 실제 데이터 세트를 사용하여 거리

장면의 자동 분류를 연구하여 기계학습을 통해 거리의 청결도 수준 및 쓰레기 물체를 식별했다. 먼저 LASAN과 협력하여 정확한 라벨(예: 부피가 큰 품목, 일반 쓰레기 불법 투기, 노숙자 야영지, 옷자란 초목, 깨끗한 거리)이 있는 22,000개의 위치 정보가 태그된 거리 이미지 데이터 세트를 사용하여 거리 청결도 수준에 기반한 이미지 분류를 시도했다(그림 3 참조). 이를 위해, 모든 이미지에서 시각적 특징 추출이 사전에 처리되었다: 즉, 색상 특징 추출, 색상 히스토그램, SIFT-BoW, CNN(Convolutional Neural Network) 등. 이러한 데이터 세트가 일회성 사전 작업으로 준비되어 공유된 후 다양한 분석을 위해 많은 기계학습 알고리즘이 적용되었다. 개별 학습은 매우 유연하고 빠른 방식으로(즉, 여러 사람이 병렬로) 통합 플랫폼 상에서 자체 데이터 세트를 독립적으로 선택할 수 있었다. 따라서 다양한 분류기(classifier)의 학습을 위해 최적의 특징을 식별하여 거리 청결 수준에 따라 트레이닝 데이터 세트를 지정할 수 있었다. 전체 지역 또는 일부 지역을 구분하여 글로벌 또는 로컬 분류기를 위한 트레이닝 데이터 세트를 간단하게 만들 수 있었다. 그림 4는 전체 데이터 세트와 지역의 하나의 데이터 세트로 사용해 실험에서 사용된 모든

1) 쓰레기의 종류에 따라 수거 장비와 방법, 폐기장의 위치가 달라지고 로스엔젤레스시와 같이 교통이 혼잡한 곳에서는 미리 적절한 준비를 하는 것이 시간과 노력을 줄이는 데 중요하다.



(그림 4) 이미지 분류 결과

기능 및 분류기와 이들의 각 조합에 대한 F1 점수를 보여줌으로써 전체 실험을 요약해 보여준다. 분류 결과는 CNN에서 제공하는 풍부한 기능 표현으로 인해 CNN 이미지 피처를 사용할 때 가장 좋은 F1 점수를 얻었음을 보여준다. 모든 분류 범주에서 0.8보다 높은 F1 점수를 획득하여 실제 사용할 수 있을 정도의 정확도를 보였다.

쓰레기 분류 모델이 학습된 후 신고되어 들어오는 새로운 이미지의 자동 식별이 이루어진다. 레이블이 지정되지 않은 새로운 이미지의 분류가 완료되면 그 결과는(즉, 부피가 큰 품목, 쓰레기 불법 투기, 노숙자 야영지, 옷자란 초목, 깨끗한 거리 중 하나의 레이블) 데이터베이스의 원본 이미지에 대한 별도의 주석으로 추가된다. 그러면 다른 이해 관계자가 플랫폼 상에서 다른 독립적 분석을 위해 데이터와 분석 결과를 공유하고 활용할 수 있다. 예를 들어, LASAN이 수집한 데이터 세트를 다른 도시에서 비슷한 응용을 위한 전이학습(transfer learning)에 활용할 수 있다. 또는 동일한 이미지 데이터 세트를 사용하여 거리의 낙서를 식별하기 위한 별도의 학습을 수행하고 데이터 세트에 별도의 결과로 주석을 달 수도 있다. 이러한 방식으로 다양한 시각적 분석을 수행할 수 있으며 그 결과를 주석으로 달고 공유하여 보다 포괄적인 시각 정보 데이터베이스로 발전하게 된다. 새로 배운 정보를 다른 분석으로 변환할 수도 있다. 실험 결과 중 하나가 거리에

서 노숙자가 야영하는 텐트를 인식하는 것이기 때문에 LASAN 데이터의 거리 청결 분류 결과를 로스엔젤레스시의 노숙자 수 산정에 활용할 수 있다. 즉, 한 정부기관의 거리 청결에 대한 분석에서 얻은 지식이 독립적인 다른 정부기관에서 별도의 데이터 수집 및 추가 학습 처리 없이 노숙자 수 계산과 같은 사회 연구에 유용한 정보로 사용하도록 변환되었다. 추가로 위치가 태그된 야영지 이미지를 면밀히 조사하여 노숙자의 삶에 대한 추가 지식을 알아낼 수도 있다.

3.2 낙서 제거

로스엔젤레스시 거리의 문제 중 하나는 벽을 포함해 온갖 곳에 그려진 수많은 낙서들이다. 이런 낙서들을 제거하기 위해 필요한 정보는 위치와 낙서가 그려진 표면 물질을 아는 것이다. 현장에 가기 전에 콘크리트, 나무, 유리 등과 같은 표면 물질을 파악하는 것이 제거 방법을 미리 결정해 준비할 수 있기에 큰 도움이 된다. 따라서, 이미지 기계학습을 통해 자동으로 표면 물질을 파악하는 연구를 실행했다[5]. 데이터 세트는 LASAN에서 수집한 19,000개의 낙서 이미지로 구성되었다. 모든 이미지는 하나 이상의 낙서가 있으며 그 주위를 박스 형태로 태그가 되었으며 18 가지 유형의 표면물질과 형태(콘크리트 담, 나무 문, 철 기둥, 등등, 그림 5 참조) 중 하나로 레이블이 되어있다. 이러한 이미지는 낙서 제거 프로세스를 자동화하기 위해 적용 소재의 분류에 따라 LASAN 전문가가 올바르게 레이블을 지정한 것이다.

기술적으로 낙서로 덮여진 물질 인식은 이미지 상에서 낙서 감지와 그 밑의 물질 인식이라는 두 가지 작업으로 구성된다. 물질 인식을 위해 하이퍼 스펙트럼 이미징 카메라(반사파 측정 기

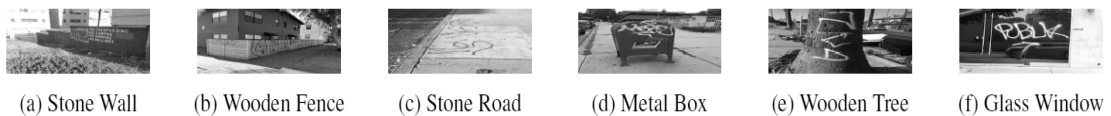
반)와 같은 특수 장치를 사용하면 낙서로 덮힌 표면 물질의 인식 작업을 정확하게 해결할 수 있으나 대규모 모니터링에서는 비용이 많이 들고 실용적이지 않다. 다른 방법으로 일반 카메라에서 사용자가 생성한 이미지를 사용해 분석할 수 있다. 그러나 기존의 일반 물질 인식을 위한 이미지 기반 접근 방법들은[12,13] 표면이 낙서로 덮여 있는 경우 비효율적일 수 있다. 이를 극복하기 위해 우리는 새로운 이미지 학습 기반 접근 방식을 제안하였다.

우리의 접근 방식은 학습을 두 단계로 나누는 것이다. 첫번째 단계는 표면 물질의 유형에 관계없이 덮여 쓴 낙서의 특성만 학습하여 낙서 부위를 찾아내는 데 초점을 맞추고, 두번째 단계는 덮여진 물질 유형을 구별되는 특성을 학습하는 다른 모델을 개발하는 것이다. 이 접근법은 2 단계 학습 접근법(그림 6 참조)이라고 하며, 낙서를 잠재적으로 포함하는 이미지 부분을 감지하는 감지 모델과 물질 분류 모델의 두 가지 모델로 구성된다. LASAN에서 제공한 데이터 세트를 이용한 실험 결과에 따르면 우리가 제안한 2 단계 학습법은 기존의 표면 물질 인식 모델과 비교해 최대 2.1배 향상된 정확도를 보여주었다. 다양하고 복잡한 물질 인식에 있어 평균적으로 0.6

mAP(man Average Precision)를 보여주었고 stone wall과 같이 많은 트레이닝 이미지가 있는 경우엔 더 높은 정확도를 나타내었다.

3.3 도로면 손상 인식

도시의 경제는 본질적으로 공공 시설과 인프라의 영향을 받고 이러한 인프라의 기본 요소는 도로이다. 그러나 많은 요인(예: 비와 노후화)이 도로 효율성, 운전자 안전 및 차량 가치에 심각한 영향을 미치는 다양한 유형의 도로 손상을 유발하고 있다. 따라서 국가는 도로 유지 보수 및 재건에 많은 연간 예산을 투자한다. 효율적인 도로 유지 관리 계획에는 안정적인 모니터링 시스템이 필요하다. 가장 간단한 방법은 육안 검사이나 비용이 많이 들고 시간이 많이 걸리기 때문에 효율적이지 못하다. 따라서 연구자들은 진동 기반, 레이저 스캔 기반 및 이미지 기반 방법을 포함하여 자동 도로 손상 검사를 위한 여러 솔루션을 개발해왔다. 진동 방법이나 레이저 스캐닝 방법은 도로 상태에 대한 정확한 정보를 제공하지만 비용이 많이 들고 특히 도로 폐쇄가 필요한 것이 문제이다. 교통이 혼잡한 로스엔젤레스에서 도로 폐쇄를 하고 모니터링한다는 것은 사회적



(그림 5) 낙서가 새겨진 표면 물질들의 예



(그림 6) 2 단계 학습 방법

〈표 1〉 도로면 손상 분류

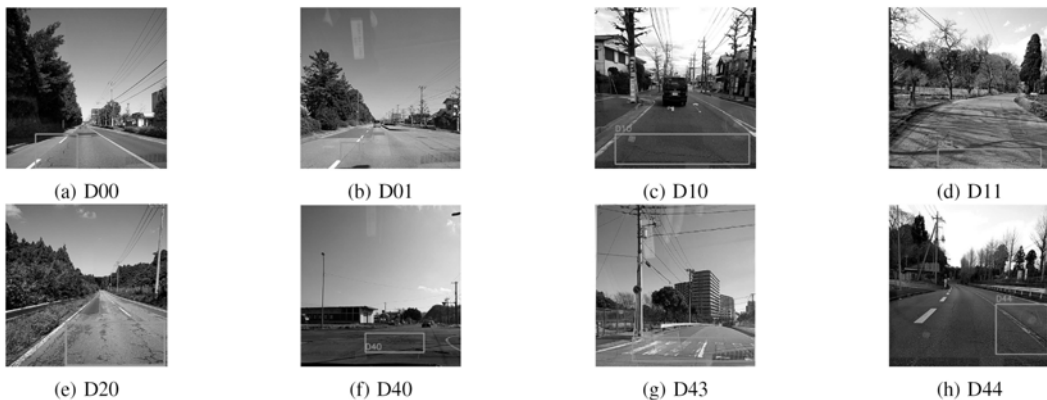
Damage Type		Detail	Class Name
Crack	Linear Crack	Longitudinal	Wheel mark part D00
			Construction joint part D01
	Lateral	Equal interval D10	
		Construction joint part D11	
	Alligator Crack	Partial pavement, overall pavement D20	
Other Corruption	Rutting, bump, pothole, separation D40		
	Crosswalk blur D43		
	White/Yellow line blur D44		

비용이 너무 높다. 한편 이미지 처리 방법은 저렴하지만 정확성이 부족한 점이 문제이다. 하지만 아직은 미성숙함에도 불구하고 최근 이미지 분석 기술의 발전은 인상적인 결과를 만들어 내고 있기에 우리는 이미지 분석을 통해 도로 표면 손상을 인식하고 분류하는 연구를 실행했다.

LASAN이 운영하는 700대가 넘는 쓰레기 수거 차량의 후면 카메라는 아래를 향해 도로면을 녹화하고 있기에 이를 이용해 도로면 손상을 파악할 수 있다. 하지만 기계학습에 필요한 대용량의 이미지 레이블링이 준비되지 않았기에 2018년 IEEE BigData Cup Road Damage Detection Challenge에서 제공한 이미지 데이터 세트(7,231 트레이닝 이미지, 1,813 테스트 이미지)를 이용해 연구를 수행했다[4]. 이 데이터는 스마트폰을

이용해 일본의 7 개 지역에 위치한 도시의 도로면을 녹화한 것으로 각 이미지는 하나 이상의 관심 영역에 손상 상태를 나타내는 레이블이 기록되어있다[11]. 표 1에서 볼 수 있듯이 도로면 손상 분류에는 균열과 기타 손상의 두 가지 범주로 일반화할 수 있는 8 가지 유형이 포함되어 있다(그림 7). 도로 손상 유형 감지 문제를 해결하기 위해 각 도로 손상 유형을 구별 가능한 클래스로 정의하고 각 도로 손상 유형의 시각적 특징 패턴을 학습하기 위해 도로 손상 데이터 세트에 대해 훈련할 최신의 물체 인식 알고리즘(즉, YOLO V3) 중 하나를 사용하였다.

일부 클래스(예: D10, D11, D40 및 D43)가 다른 클래스에 비해 이미지 수가 적어 도로 손상 클래스 간의 데이터 세트 분포가 균형을 이루지



(그림 7) 도로면 손상 클래스별 이미지 예

않기에 Python Augmentor 라이브러리를 사용하여 이러한 클래스를 포함하는 트레이닝 데이터 세트를 위한 합성 이미지를 생성하여 사용하였다. 합성 시 도로 손상 장면이 영향을 받지 않도록 하기 위해 Augmentor 도구에서 제공하는 이미지 처리 기술(예: 밝기, 회색조)을 신중하게 선택하여 별도의 증강 데이터 세트를 만들어 학습에 사용하였다. 트레이닝 데이터 세트를 처리할 때 고려할 또 다른 기술은 크롭핑(cropping)이다. 잘린 이미지에 레이블이 달린 영역이 포함되도록 모든 이미지를 잘라서 원본 이미지 보다 더 작은 크기의 데이터 세트를 만들었다. 이러한 데이터 세트를 사용하면 학습모델이 학습(도로면 손상)과 관련 없는 부분(예: 도로 위 건물이나 하늘)을 삭제하여 관심 영역의 기능을 적절하게 학습하는데 집중할 수 있다. 결과적으로 원본(D), 증강(Da) 및 크롭(Dc)의 세 가지 트레이닝 데이터 세트를 준비해 학습에 사용하였다.

우리의 솔루션은 YOLO를 사용하여 모델을 학습하여 분석된 이미지에서 다양한 유형의 도로 손상을 식별 가능한 물체로 감지한다. 객체 감지기를 만들기 위해 YOLO 프레임 워크(버전 3)를 사용하여 darknet53 모델을 미세 조정하였다. 솔루션을 철저히 실험하기 위해 매개 변수들을 변경하며 각 데이터 세트(D, Da 및 Dc)를 사용하여 학습된 모델의 다양한 버전을 생성했다. D, Da 및 Dc를 사용하여 훈련된 도로 손상 인식 모델은 최대 0.61, 0.62 및 0.60의 F1 점수를 달성하였다.

4. 결 론

이 논문에서는 사용자가 효율적이고 민주적인 방식으로 시각데이터의 수집, 관리, 분석 및 공유

를 수행할 수 있게 하는 시각데이터 통합 플랫폼과 로스엔젤레스시의 시각데이터 응용 사례들을 소개하였다. 실제 데이터 세트를 사용한 실시간 서비스 구현 사례는 아직 수행할 작업이 많이 남아 있지만 제안된 핵심 개념의 실행 가능성을 보여주는 사례연구들이 소개되었다. 이 연구의 장점은 적절하게 설계된 시각데이터 플랫폼이 현재 상당히 단편화 되어있는 시각데이터의 수집, 관리 및 분석을 구조화, 통합 및 조직화하여 스마트시티 애플리케이션 설계를 용이하게 할 것이라는 가설을 평가하는 데 있다. 대규모 스마트시티 프로젝트들의 여러 이해 관계자를 고려하여 사회 문제에 대한 보다 효율적이고, 공유 가능하며, 경제적이며, 지속 가능한 시각데이터 중심의 민주적 솔루션을 만들 수 있음을 보여주었다.

참 고 문 헌

- [1] A. Alfarrarjeh, S. H. Kim, S. Agrawal, M. Ashok, S. Y. Kim, and C. Shahabi, "Image classification to determine the level of street cleanliness: A case study," in BigMM. IEEE, 2018.
- [2] S. H. Kim, J. Shi, A. Alfarrarjeh, D. Xu, Y. Tan, and C. Shahabi, "Real-time traffic video analysis using intel viewmont coprocessor," in DNIS. Springer, 2013, pp. 150-160.
- [3] A. Alfarrarjeh, S. Agrawal, S. H. Kim, and C. Shahabi, "Geospatial multimedia sentiment analysis in disasters," in DSAA. IEEE, 2017, pp. 193-202.
- [4] A. Alfarrarjeh, D. Trivedi, S. H. Kim, and C. Shahabi, "A deep learning approach for road damage detection from smartphone images," in Big Data. IEEE, 2018, pp.

- 5184-5187.
- [5] A. Alfarrarjeh, D. Trivedi, S. H. Kim, H. Park, C. Huang, and C. Shahabi, "Recognizing material of a covered object: A case study with graffiti," IEEE International Conference on Image Processing (ICIP), 2019.
- [6] S. H. Kim, Y. Lu, J. Shi, A. Alfarrarjeh, C. Shahabi, G. Wang, and R. Zimmermann, "Key Frame Selection Algorithms for Automatic Generation of Panoramic Images from Crowdsourced Geo-tagged Videos," in W2GIS. Springer, 2014, pp. 67-84.
- [7] G. Wang, Y. Lu, L. Zhang, A. Alfarrarjeh, R. Zimmermann, S. H. Kim, and C. Shahabi, "Active Key Frame Selection for 3d Model Reconstruction from Crowdsourced Geo-tagged Videos," in ICME. IEEE, 2014, pp. 1-6.
- [8] "MyLA311 App," <https://www.lacity.org/myla311>.
- [9] Q. Contributor, "Why is it important to democratize machine learning?" 2016. [Online]. <https://www.forbes.com/sites/quora/2016/12/28/why-is-it-important-to-democratize-machine-learning/#39ca0f17582a>
- [10] D. Wellers, J. Woods, D. Jendroska, and C. Koch, "Why machine learning and why now?" 2017. [Online]. <http://www.digitalistmag.com/executive-research/why-machine-learning-and-why-now>
- [11] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiya, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," CACAIE.
- [12] S. Bell, P. Upchurch, N. Snavely, et al., "Material recognition in the wild with the

materials in context database," in CVPR, 2015, pp. 3479-3487.

- [13] P. Wieschollek and H. Lensch, "Transfer learning for material classification using convolutional networks," arXiv preprint arXiv:1609.06188, 2016.
- [14] S. H. Kim, A. Alfarrarjeh, G. Constantinou, C. Shahabi. "TVDP: Translational Visual Data Platform for Smart Cities". International Workshop on Data-Driven Smart Cities in conjunction with ICDE, April 2019.

저 자 약 력



김 선 호

이메일 : seonkim@usc.edu

- 1986년 연세대학교 전자공학과 (학사)
- 1994년 University of Southern California 컴퓨터공학과 (석사)
- 1999년 University of Southern California 전산학과 (박사)
- 1999년~2008년 University of Denver 전산학과 교수
- 2008년~2010년 University of District of Columbia 전산학과 교수
- 2010년~현재 University of Southern California, IMSC(Integrated Media Systems Center) 연구소 부소장
- 관심분야: 멀티미디어 응용, 데이터과학, 이미지 기계학습, 스마트시티