

# 건설현장 영상 분석을 위한 웹 크롤링 기반 학습 데이터베이스 구축 자동화

황정빈\* · 김진우\*\* · 지석호\*\*\* · 서준오\*\*\*\*

Hwang, Jeongbin\*, Kim, Jinwoo\*\*, Chi, Seokho\*\*\*, Seo, JoonOh\*\*\*\*

## Automated Training Database Development through Image Web Crawling for Construction Site Monitoring

### ABSTRACT

Many researchers have developed a series of vision-based technologies to monitor construction sites automatically. To achieve high performance of vision-based technologies, it is essential to build a large amount and high quality of training image database (DB). To do that, researchers usually visit construction sites, install cameras at the jobsites, and collect images for training DB. However, such human and site-dependent approach requires a huge amount of time and costs, and it would be difficult to represent a range of characteristics of different construction sites and resources. To address these problems, this paper proposes a framework that automatically constructs a training image DB using web crawling techniques. For the validation, the authors conducted two different experiments with the automatically generated DB: construction work type classification and equipment classification. The results showed that the method could successfully build the training image DB for the two classification problems, and the findings of this study can be used to reduce the time and efforts for developing a vision-based technology on construction sites.

**Key words :** Construction sites, Vision-based analytics, Web crawling, Training database, Construction work type, Construction equipment

### 초 록

건설현장 영상 자동 모니터링을 목적으로 많은 연구자들이 영상분석기술을 활발히 개발하고 있다. 높은 성능의 영상분석기술을 개발하기 위해서는 다량의 고품질 학습용 이미지 데이터베이스(Database, DB)를 구축해야 한다. 하지만 기존의 학습 DB 구축 방법은 사람이 건설현장을 직접 방문하여 카메라를 설치하고 각각의 목적에 알맞은 영상을 수집하여 DB를 직접 구축하고 있기 때문에 이에 많은 비용과 시간이 요구된다. 뿐만 아니라 이 같은 사람 의존적인 방식은 건설현장의 다양한 특성을 모두 반영한 DB를 구축하는 것에 한계가 있다. 이러한 한계를 극복하기 위해서 본 연구는 웹 크롤링(Web Crawling) 기법을 활용하여 건설현장 영상분석을 위한 학습 이미지 DB를 자동으로 구축하는 프레임워크를 제안한다. 프레임워크 검증을 위해 건설공종과 건설장비에 대한 학습 DB를 구축하여 영상분석모델을 학습 및 평가하는 실험을 진행하였다. 그 결과, 건설현장 모니터링을 위한 학습용 이미지 DB를 자동으로 구축할 수 있었을 뿐만 아니라 이를 토대로 개발한 영상분석모델이 건설공종과 건설장비를 성공적으로 분류하는 것을 확인하였다. 결과적으로 기존의 방식보다 학습 DB를 구축하는 데 필요한 시간과 비용을 최소화할 수 있었다.

**검색어 :** 건설현장, 영상 분석, 웹 크롤링, 학습 데이터베이스, 건설공종, 건설장비

\* 서울대학교 건설환경공학부 석박통합과정 (Seoul National University · [jb.hwang@hotmail.com](mailto:jb.hwang@hotmail.com))

\*\* 정희원 · 서울대학교 건설환경종합연구소 선임연구원 (Seoul National University · [jinwoo92@snu.ac.kr](mailto:jinwoo92@snu.ac.kr))

\*\*\* 종신희원 · 교신저자 · 서울대학교 건설환경공학부 부교수 (Corresponding Author · Seoul National University · [shchi@snu.ac.kr](mailto:shchi@snu.ac.kr))

\*\*\*\* 홍콩이공대학 조교수 (Hong Kong Polytechnic University · [joonoh.seo@polyu.edu.hk](mailto:joonoh.seo@polyu.edu.hk))

Received September 10, 2019/ revised September 20, 2019/ accepted September 27, 2019



들의 위치, 동선 및 행동을 파악하여 작업 수행시 안전상태를 평가하였으며, Luo et al.(2018)은 Convolutional Neural Network와 Bayesian Learning을 활용하여 근로자들의 행동과 자세 등을 파악해 작업유형을 자동으로 분석하는 연구를 진행하였다. Golparvar-Fard et al.(2013)은 굴삭기와 덤프트럭의 Spatio-temporal (시공간적 특성)을 분석하여 시간에 따른 장비의 행동을 자동 인식하였다. Kim et al.(2018)은 건설현장 영상 데이터에서 장비를 인식한 후 LSTM 모델을 활용해 장비 간의 상호작용을 분석하여 작업유형을 판단하는 프레임워크를 개발하였다. Cai et al.(2019) 또한 LSTM 모델을 활용하여 건설장비와 작업자의 자세, 운동방향, 위치 등을 학습해 건설요소 간 상호작용을 분석함으로써 작업유형을 판단하는 프레임워크를 개발하였다.

이처럼 많은 연구들이 진행되었으나 기존의 방법들은 여전히 학습 DB를 구축하기 위해서 많은 비용과 시간을 요구하고, 이로 인해 건설현장 자동 모니터링을 위한 영상분석기술을 개발하는데 어려움이 있다. 따라서 본 연구는 웹 크롤링 기법을 활용하여 건설현장 자동 모니터링을 위한 학습용 이미지 DB를 자동으로 구축하는 프레임워크를 제안하여 학습 DB 구축에 드는 비용과 시간을 단축시키고자 한다.

### 3. 웹 크롤링 기반 학습 데이터베이스 자동 구축 프레임워크

본 연구에서 개발하는 웹 크롤링 기반 학습용 이미지 DB 자동 구축 프레임워크는 Fig. 2와 같이 총 3단계로 구성된다. 먼저, (1) 건설현장에서 사용되는 공중 및 장비들을 인터넷에서 효율적으로 검색하기 위해서 건설공사표준분류체계를 토대로 건설현장 맞춤형 검색 키워드 사전을 구축한다. 그 다음, 구축한 검색 키워드 사전을 바탕으로 (2) 웹 크롤링 기법을 이용하여 인터넷 웹사이트 (Google, Bing 등)에 산재한 다량의 건설현장 이미지 데이터를 자동으로 검색하고 수집하여 학습용 이미지 DB를 구축한다. (3) 최종적으로 구축한 학습 DB를 통해 딥러닝 기반 영상분석모델을 개발한다. 각 단계별 자세한 내용은 다음과 같다.

#### 3.1 검색 키워드 사전 구축

인터넷 검색 시 원하는 데이터를 효과적으로 추출해내기 위해 MasterFormat, UniFormat 등 대표적인 건설공사표준분류체계를 토대로 이미지 데이터 검색에 최적화된 키워드 사전을 Tables 1 and 2와 같이 구축하였고, 건설공종 및 건설장비 검색 시 구축된 키워드를 사용하여 검색을 진행하였다. 예를 들면, 건설공종의



Fig. 2. Framework for Automated Training DB Development through Image Web Crawling

Table 1. Keyword Dictionary for Construction Work Type

Earthwork	Framing	Finishing
Earthwork	Reinforcement construction	Curtain wall installation
Earthwork construction	Framing construction	Curtain wall progress
Earthmoving construction	Frame construction	Curtain wall worker

Table 2. Keyword Dictionary for Construction Equipment

Crane	Dump Truck	Excavator	Mixer Truck	Roller
Mammoet crane	Ashok Layland dump truck	Caterpillar excavator	Hyundai mixer truck	ABG road roller
Sarens crane	Ford dump truck	Komatsu excavator	Hanson mixer truck	Caterpillar road roller
Lampson crane	GMC dump truck	Volvo excavator	Liebherr mixer truck	Hitachi road roller
ALE crane	Mercedes Benz dump truck	Kubota excavator	Mcneilus mixer truck	Bomag road roller
Prangl crane	Chevy dump truck	JCB excavator	Aimix mixer truck	Volvo road roller

경우 토공사 검색 시 ‘earthwork’, ‘earthwork construction’, ‘earthmoving construction’ 등과 같은 키워드를 사용하였다. 건설장비의 경우 다양한 제조사의 모델에 대한 이미지 데이터를 수집하기 위해서 ‘Volvo excavator’, ‘Kubota excavator’, ‘Komatsu excavator’, ‘JCB excavator’, ‘Caterpillar excavator’ 등과 같이 제조사명을 키워드에 포함하였다. 이를 통해 웹사이트에서 다량의 다양한 이미지 데이터를 검색할 수 있고, 최종적으로 양질의 학습용 이미지 DB 구축이 가능하다.

### 3.2 웹 크롤링 및 학습 DB 구축

웹 크롤링 및 학습 DB 구축 단계는 앞서 개발된 검색 키워드 사전을 토대로 컴퓨터가 인터넷 웹사이트에서 이미지 데이터를 자동으로 검색하고 저장하는 역할을 수행한다. 저장된 이미지 데이터는 검색 키워드에 따라 분류되며, 이를 통해 학습용 DB를 자동으로 구축할 수 있다. 본 방법은 기존의 사람 의존적인 학습 DB 구축 방법과 비교할 때 보다 빠른 속도로 양질의 DB를 자동으로 구축할 수 있으며, 이를 구현하기 위해서 Python 3.7 환경에서 request, urllib, BeautifulSoup, selenium 등의 패키지를 활용하였다.

### 3.3 영상분석모델 개발

본 단계의 목적은 앞서 구축된 학습용 이미지 DB를 이용하여

영상분석모델을 개발하는 것이다. 특히, 본 연구에서는 ImageNet DB의 1,000,000개 이상의 이미지 데이터로 학습된 48개의 계층을 가진 Convolutional Neural Network인 inception\_v3 딥러닝 모델을 건설공종 및 장비 분류에 맞춤 개발하였다. 해당 모델을 학습하기 위해서 Stochastic Gradient Descent 최적화 알고리즘을 이용하였고 1회당 100번의 학습을 진행하였다. 해당 학습 모델의 Hyperparameter는 TensorFlow에서 제공하는 값들인 batch size는 32, learning rate은 초기값 0.01, 최종값은 0.0001을 활용하였다.

## 4. 실험 결과 및 분석

본 연구에서 제안한 프레임워크를 검증하기 위해서 두 가지 실험을 수행하였다. 첫 번째로 웹 크롤링을 통해 구축한 학습용 건설공종 DB를 이용하여 공종을 자동으로 분류하였고, 두 번째는 학습용 건설장비 DB를 통해 장비의 종류를 분류하였다.

### 4.1 실험 1: 건설공종 DB 구축 및 모델 개발

첫 번째 실험으로 건설공종을 분류하는 영상분석모델을 개발하였다. 건설공종은 대표적인 건축공종인 토공사, 골조공사, 마감공사 총 3가지를 분류 대상으로 선정하였고, 각 공종의 검색 키워드는 3개씩으로 Table 1과 같이 선별하였다. 웹 크롤링 기법을 활용한

Table 3. Example of Construction Work Type Classification

Example images of construction work type			
Predicted Class & Confidence	Earthwork (94.23 %)	Framing (88.97 %)	Finishing (83.14 %)

Table 4. Example of Construction Work Type False Classification

Example images of construction work type		
Actual Class	Framing	Finishing
Predicted Class	Finishing	Framing

검색 및 학습 DB 구축 결과, 구축된 이미지 데이터는 토공사 3,862장, 골조공사 3,726장, 마감공사 3,693장으로 총 11,281장이며 이 중 7,897장(70 %)은 학습 데이터, 3,384장(30 %)은 시험 데이터로 활용하였다. 개발한 영상분석모델을 적용한 공종분류 결과 예시는 Table 3과 같으며, 자동으로 구축된 DB를 활용하여서도 건설공종을 성공적으로 분류하는 것을 볼 수 있다. 평균 정확도는 80.3 %로 나타났고, 공종별 정확도는 Fig. 3과 같이 토공사 88 %, 골조공사 80 %, 마감공사 70 % 순으로 도출되었다. 토공사의 경우, 다른 공종과 비교하여 고품질의 이미지 데이터를 손쉽게 수집할 수 있었기 때문에 학습이 용이하였고 타 공종에 비해 정확도가 높아진 것으로 분석된다. 또한, 골조공사와 마감공사에는 구조물, 작업자 등 투입되는 건설자원이 비슷하여 그 시각적 특성이 유사하

기 때문에 Table 4와 같이 영상분석모델이 두 가지 공종을 오분류하는 경우도 있었다.

#### 4.2 실험 2: 건설장비 DB 구축 및 모델 개발

두 번째 실험으로 대표적인 건설장비인 크레인, 덤프트럭, 굴삭기, 레미콘, 롤러를 자동으로 분류하는 영상분석모델을 개발하였다. 이에 따라 웹 크롤링 기법을 활용하여 Table 2와 같이 장비 유형별 5개의 검색 키워드로 이미지 데이터를 자동 수집하였다. 검색 및 학습 DB 구축 결과, 구축된 이미지 데이터는 크레인 2,787장, 덤프트럭 3,274장, 굴삭기 3,273장, 레미콘 3,170장, 롤러 3,014장으로 총 15,518장이며 실험 1과 동일한 비율로 10,863장(70 %)은 학습 데이터, 4,655장(30 %)은 시험 데이터로 활용하였다. 개발한

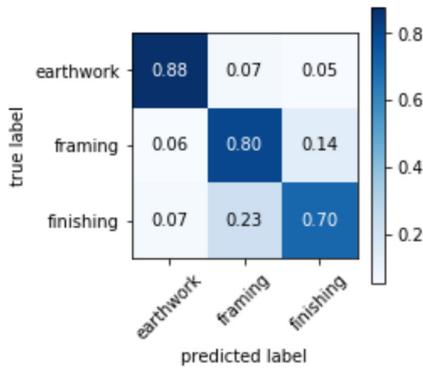


Fig. 3. Confusion Matrix of Construction Work Type Classification Result

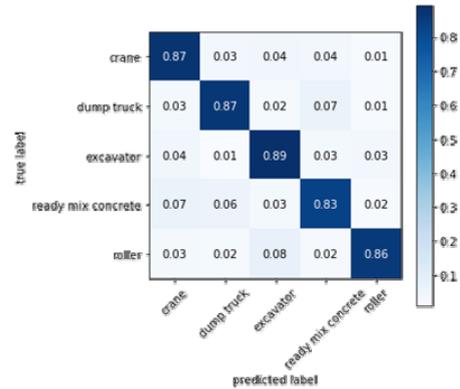


Fig. 4. Confusion Matrix of Construction Equipment Classification Result

Table 5. Example of Construction Equipment Classification

Example images of construction equipment					
Predicted Class & Confidence	Crane (95.30 %)	Dump Truck (95.03 %)	Excavator (96.21 %)	Mixer Truck (95.67 %)	Roller (96.56 %)

Table 6. Example of Construction Equipment False Classification

Example images of construction equipment		
Actual Class	Dump Truck	Mixer Truck
Predicted Class	Mixer Truck	Dump Truck

영상분석모델을 적용한 결과 예시는 Table 5와 같다. 평균 정확도는 86.4 %로 나타났고, 장비별 정확도는 Fig. 4와 같이 크레인 87 %, 덤프트럭 87 %, 굴삭기 89 %, 레미콘 83 %, 롤러 86 % 순서로 도출되었다. 대체적으로 건설장비는 기하학적 형태 등 각각의 특성이 뚜렷하기에 건설공종 분류(실험 1) 대비 정확도가 높은 것으로 분석된다. 다만, 덤프트럭과 레미콘의 경우 시각적 특성이 유사하여 Table 6과 같은 혼동이 발생하는 것으로 파악된다.

## 5. 결론

본 연구는 건설현장 자동 모니터링 및 영상분석기술 개발에 필요한 학습용 DB를 웹 크롤링 기법을 활용하여 자동으로 구축하는 방법을 제안하였다. 이를 검증하기 위해서 실제 건설공종과 건설장비를 분류하는 두 가지 실험을 진행하였으며, 그 결과 10,000장의 학습용 이미지 DB를 구축하는 데 단 30분 가량의 시간만이 소요되었다(1개 이미지 데이터당 약 0.18초). 이는 사람이 일일이 인터넷에서 사진들을 검색하여 각 사진을 컴퓨터에 저장한 후 학습용 이미지 데이터를 Labeling하는 데 걸리는 시간(1개 이미지 데이터당 약 90초)과 비교할 때 약 500배 빠른 것으로 파악되었다. 또한, 영상분석모델의 성능을 살펴보면, 공종분류의 정확도는 80.3 %, 장비분류의 경우 정확도 86.4 %로, 신뢰할 수 있는 수준의 정확도를 보였다. 결과적으로 본 연구 결과를 활용하여 웹사이트에 산재한 다양한 이미지 데이터를 자동 수집함으로써, 학습 DB 구축 및 영상분석기술 개발을 위한 시간과 비용을 최소화할 수 있을 뿐만 아니라 건설현장을 자동으로 모니터링하는 데 기여할 수 있다.

## 감사의 글

본 연구는 국토교통부 국토교통기술촉진연구사업의 연구비지원(19CTAP-C151784-01)에 의해 수행되었습니다.

## References

- Ahuja, M. S., Singh, J. and Varnica, B. (2014). "Web crawler: Extracting the web data." *International Journal of Computer Trends and Technology*, Vol. 13, No. 3, pp. 132-137.
- Arabi, S., Haghghat, A. and Sharma, A. (2019). *A deep learning based solution for construction equipment detection: from development to deployment*. arXiv preprint arXiv:1904.09021.
- Cai, J., Zhang, Y. and Cai, H. (2019). "Two-step long short-term memory method for identifying construction activities through positional and attentional cues." *Automation in Construction*, Vol. 106, p. 102886.
- Chi, S. H. and Caldas, C. H. (2011). "Image-Based safety assessment: Automated spatial safety risk identification of earthmoving and surface mining activities." *Journal of Construction Engineering and Management*, Vol. 138, No. 3, pp. 341-351.
- Chi, S. H., Caldas, C. H. and Kim, D. Y. (2009). "A methodology for object identification and tracking in construction based on spatial modeling and image matching techniques." *Computer-Aided Civil and Infrastructure Engineering*, Vol. 24, No. 3, pp. 199-211.
- Fard, M. G., Heydarian, A. and Niebles, J. C. (2013). "Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers." *Advanced Engineering Informatics*, Vol. 27, No. 4, pp. 652-663.
- Gong, J. and Caldas, C. H. (2010). "Computer vision-based video interpretation model for automated productivity analysis of construction operations." *Journal of Computing in Civil Engineering*, Vol. 24, No. 3, pp. 252-263.
- Kim, J. W., Chi, S. H. and Seo, J. W. (2018). "Interaction analysis for vision-based activity identification of earthmoving excavators and dump trucks." *Automation in Construction*, Vol. 87, pp. 297-308.
- Li, H., Lu, M., Hsu, S. C., Gray, M. and Huang, T. (2015). "Proactive behavior-based safety management for construction safety improvement." *Safety Science*, Vol. 75, pp. 107-117.
- Luo, X., Li, H., Yang, X., Yu, Y. and Cao, D. (2018). "Capturing and understanding workers' activities in far-field surveillance videos with deep action recognition and bayesian nonparametric learning." *Computer-Aided Civil and Infrastructure Engineering*, Vol. 34, pp. 333-351.
- Rezazadeh Azar, E. and McCabe, B. (2011). "Automated visual recognition of dump trucks in construction videos." *Journal of Computing in Civil Engineering*, Vol. 26, No. 6, pp. 769-781.
- Seo, J. O., Han, S. V., Lee, S. H. and Kim, H. K. (2015). "Computer vision techniques for construction safety and health monitoring." *Advanced Engineering Informatics*, Vol. 29, No. 2, pp. 239-251.
- Yang, J., Vela, P., Teizer, J. and Shi, Z. (2012). "Vision-Based tower crane tracking for understanding construction activity." *Journal of Computing in Civil Engineering*, Vol. 28, No. 1, pp. 103-112.