

https://doi.org/10.7236/JIIBC.2021.21.6.111
JIIBC 2021-6-16

딥러닝 학습을 위한 초분광 영상 데이터 관리 소프트웨어 개발

Management Software Development of Hyper Spectral Image Data for Deep Learning Training

이다빈*, 김홍락*, 박진호*, 황선정*, 신정섭**

Da-Been Lee*, Hong-Rak Kim*, Jin-Ho Park*,
Seon-Jeong Hwang*, Jeong-Seop Shin**

요약 초분광 영상은 적외선 영역의 전자기파 대역을 수백 개의 파장으로 나누어 영상화한 데이터로 다양한 분야에서 물체를 찾거나 분류하는 것에 활용된다. 최근에는 딥러닝을 사용하여 분류하는 방법이 주목받고 있지만 초분광 영상 데이터의 특성으로 인해 초분광 영상을 학습 데이터로 사용하기 위해서는 기존의 가시광 영상과는 다른 처리 기법이 필요하다. 이를 위해 초분광 큐브에서 특정 파장의 영상을 선택하여 Ground Truth 작업을 수행하고 환경정보를 포함하여 데이터를 관리하는 소프트웨어를 개발하였다. 본 논문에서는 해당 소프트웨어의 구성과 기능에 대하여 설명한다.

Abstract The hyper-spectral image is data obtained by dividing the electromagnetic wave band in the infrared region into hundreds of wavelengths. It is used to find or classify objects in various fields. Recently, deep learning classification method has been attracting attention. In order to use hyper-spectral image data as deep learning training data, a processing technique is required compared to conventional visible light image data. To solve this problem, we developed a software that selects specific wavelength images from the hyper-spectral data cube and performs the ground truth task. We also developed software to manage data including environmental information. This paper describes the configuration and function of the software.

Key Words : Hyper-Spectral Image, Deep Learning, Dataset Training, Data Management

1. 서 론

초분광 영상은 가시광선 영역에서 적외선 영역까지의 전자기파 대역을 수백 개의 파장으로 나누어 영상화한 데이터를 의미한다. 나누어진 파장 별 영상들을 효율적으로 관리하기 위해 각 영상 데이터를 합쳐 데이터 큐브

형태로 구성하고 이를 초분광 큐브라고 한다. 초분광 영상 데이터는 물체의 색, 형태 정보만 획득 가능한 가시광 영상과는 다르게 물체의 분광 정보를 분석하여 물체의 재질 분석이 가능하다는 장점이 있다. 이러한 특성을 이용하여 국방, 의료, 안전, 농업 등 다양한 분야에서 물체를 찾거나 분류하는 연구에 활용된다^[1, 2, 3, 8]. 예를 들어

*정회원, LIG넥스원(주)

**비회원, 국방과학연구소

접수일자 2021년 10월 28일, 수정완료 2021년 11월 28일
게재확정일자 2021년 12월 10일

Received: 28 October, 2021 / Revised: 28 November, 2021 /

Accepted: 10 December, 2021

*Corresponding Author: dabeen.lee@lignex1.com

Dept. of RF & IIR Seeker R&D Lab, LIG Nex1 Co., Ltd. Korea

농산물 종자의 발아 여부를 확인하는 연구^{4, 9, 10}와 약품 계열에 따라 분광 스펙트럼이 상이하게 나타나는 점을 이용한 초분광 영상을 활용한 약품 분류 연구가 있다⁵.

초분광 영상을 활용한 분류 기법은 전통적인 영상처리 기법부터 딥러닝까지 다양한 방법들이 있다. 최근에는 초분광 영상 데이터를 학습 데이터로 사용한 딥러닝이 좋은 결과를 보인다⁶. 딥러닝 학습을 위해서는 영상 내 목표 물체에 대한 Ground Truth 작업이 완료된 대용량의 데이터 셋이 필요하다. 초분광 영상은 가시광 영상과 다르게 데이터 형태가 큐브 형태이고 물체의 온도 및 대기 온도에 따라 큰 영향을 받는 특성을 가진다⁷. 따라서 기존의 가시광 영상 데이터 셋 생성에 사용한 Ground Truth 틀과 데이터 관리 기법을 초분광 영상 데이터 셋 생성에 그대로 사용하는 것은 부적절하다. 이를 해결하기 위해 초분광 큐브 데이터를 입력으로 사용하는 Ground Truth 틀과 환경정보를 포함하는 데이터 관리 소프트웨어를 개발하였다. 본 논문에서는 해당 소프트웨어의 구성과 기능에 대하여 설명한다.

II. 본 문

1. 초분광 영상 데이터의 특징을 반영한 기능

앞서 설명한 것과 같이 초분광 영상 데이터는 구조와 특징이 가시광 영상과 다르다. 따라서 초분광 영상 데이터를 딥러닝 학습 데이터 셋으로 사용하고자 한다면 기존의 가시광 영상 데이터 관리 소프트웨어에 추가적인 기능이 필요하다.

첫 번째로 Ground Truth 작업 시 작업자가 파장별 영상을 선택하여 작업하는 기능이 필요하다. 초분광 큐브 데이터는 파장 영역을 나누어 영상화한 데이터의 집합인데 기상 환경, 표적 물체의 종류 등에 따라 파장 영역별로 영상화된 품질이 다르다. 예를 들어 강우 환경에서 획득한 데이터의 경우 대기 환경에 영향을 크게 받는 MWIR(mid wave infrared) 대역의 영상은 영상 내 노이즈가 발생하여 품질이 좋지 못하지만 물체에 반사된 에너지를 사용하는 SWIR(short wave infrared) 대역은 상대적으로 좋은 품질의 영상을 획득할 수 있다. 이와 비슷하게 물체의 재질에 따라 특정 대역의 영상에서 영상 내 물체의 가시성이 상대적으로 더 좋아져 해당 물체의 위치 정보를 더 정확하게 획득할 수도 있다. 이러한 특징을 고려하여 초분광 영상 데이터에 Ground Truth 작업을 수행할 경우 가시광 영상에서 작업하는 것과는 달리

영상의 품질이 좋은 대역의 데이터를 선택하여 작업할 수 있도록 하는 것이 신뢰성 높은 Ground Truth 획득을 위해 필요하다.

2. 관리 데이터 종류 및 전처리 방법

본 장에서는 개발한 소프트웨어에서 관리하는 초분광 영상 데이터와 환경정보들의 전처리 및 획득 방법에 대하여 설명한다. 또한 데이터 관리의 편의를 위해 초분광 영상 데이터와 환경정보들을 합쳐 생성한 Bon 파일에 대하여 설명한다.

가. 초분광 영상 데이터

초분광 영상 데이터는 데이터 획득에 사용한 장비의 제조사들에 따라 각각 다른 저장 형식들이 존재한다. 일부 제조사들은 저장 형식의 내부 구조에 대해 공개하지 않았다. 또한 개발한 소프트웨어에서 모든 제조사들의 저장 형식들에 맞추는 것은 어렵기 때문에 데이터 관리 소프트웨어에 입력되는 데이터가 일반화되도록 획득한 초분광 영상 데이터에 전처리 과정을 수행하였다. 전처리 과정에서 초분광 큐브 데이터는 binary 파일로 따로 저장하였고 추가 헤더로 사용될 장비명, 획득 시간 등의 필요 정보들을 선별하여 csv 형식으로 저장하였다.

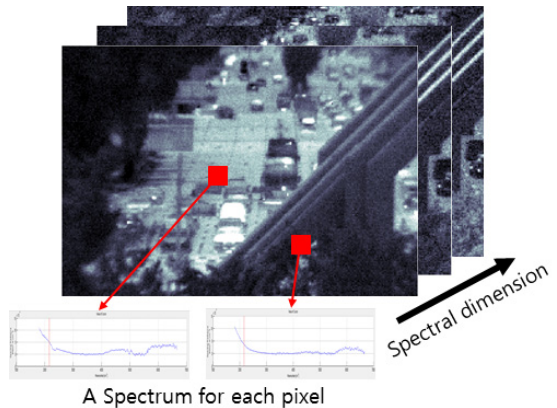


그림 1. 초분광 영상 데이터
Fig. 1. Hyper-Spectral Image Data

나. 환경정보

환경정보로는 미세먼지, 대기 온도, 강수량, 운량, 태양고도, 촬영 장소, 표적 온도를 획득하였다. 이 중 미세먼지, 대기 온도, 표적 온도는 보유하고 있는 센서를 사용하여 데이터를 획득하였고 획득한 데이터를 csv 또는 txt 형식으로 저장하여 데이터 관리 소프트웨어에서 입

력받을 수 있도록 하였다. 강수량, 운량, 태양고도와 같은 데이터는 정부 기관에서 관리하는 기상청, 생활천문관 사이트에 데이터 획득을 수행한 장소, 시간 값을 입력하고 적절한 정보를 획득하였다. 마찬가지로 csv 또는 txt 형식으로 저장하여 데이터 관리 소프트웨어에서 입력되도록 하였다. 입력된 환경정보는 대부분 그대로 사용되지만 날씨 데이터를 생성하기 위해 입력받은 강수량과 운량 정보는 데이터 값의 크기에 따라 정량적인 날씨 정보로 변환하여 데이터베이스에 저장한다.

다. Bon 파일

Bon 파일은 초분광 영상 데이터를 기준으로 해당하는 Ground Truth 라벨 그리고 환경정보들을 합친 최종 데이터 형식이다. 제안하는 데이터 관리 소프트웨어는 초분광 영상 데이터들과 1:1 대응하는 Bon 파일들을 생성하고 데이터베이스에서 조건에 따라 검색하여 원하는 Bon 파일들을 식별하여 해당 파일들을 획득할 수 있도록 한다.

Bon 파일의 구조는 그림2와 같이 영상 크기, 영상 획득 장비명 등 영상 데이터 관련 헤더, 일정한 크기의 환경정보들, 가변적 크기를 가지는 환경정보들 수, 가변 크기의 환경정보들, 초분광 큐브 데이터 순으로 구성된다. 초분광 영상 획득 장비와 영상 내 표적 수에 따라 가변적으로 변하는 정보들과 그렇지 않은 정보를 구분하여 저장공간을 효율적으로 사용할 수 있도록 하였다.



그림 2. Bon 파일 구조
 Fig. 2. Bon File Structure

3. 데이터 관리 소프트웨어 구성

앞서 설명한 데이터들을 딥러닝 학습 데이터로 사용할 수 있도록 관리하는 소프트웨어를 개발하였다. 개발된 소프트웨어는 총 3개의 모듈로 이루어져 있다. 초분광 큐브와 환경정보를 합친 최종 데이터 형태인 Bon 파일

을 생성하는 모듈, 초분광 큐브에서 물체의 위치 정보를 획득하는 Ground Truth 툴 모듈 마지막으로 데이터를 데이터베이스에 업로드, 다운로드하고 원하는 조건의 영상을 검색하여 관리하는 데이터 관리 모듈로 구성된다. 이러한 모듈들을 통해 Ground Truth 작업 후 Bon 파일을 생성하고 최종적으로 원하는 조건에 맞추어 데이터 관리가 가능하도록 하였다. 본 프로그램은 C#언어로 개발되었고 MS Access DB 기능을 이용하였다. 그림3은 데이터 처리 시 실행되는 과정을 나타낸 것이며 관련된 설명은 3.1~3.3절에서 자세히 기술하였다.

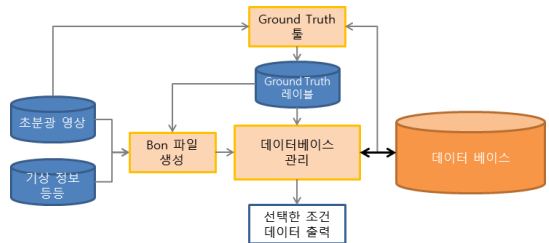


그림 3. 데이터 처리 과정
 Fig. 3. Data Processing Sequence

가. Bon 파일 생성

초분광 영상 데이터의 획득 시간을 기준으로 가장 가까운 시간에 획득한 환경정보들을 선택하여 Bon 파일을 생성한다. 정량적인 데이터인 미세먼지, 대기 온도, 태양고도, 표적 온도 등은 입력받은 값을 그대로 사용한다. 정성적인 날씨 데이터를 생성하기 위해 입력받은 강수량, 운량 데이터는 강수량 값이 0보다 큰 경우 날씨 데이터에 '비'를 입력하고 운량의 경우 운량 크기에 따라 맑음(0~2), 구름조금(3~5), 구름많음(6~8), 흐림(9~10이상)으로 날씨 데이터를 저장한다. 운량 데이터 값을 정성적인 날씨 값으로 바꾼 기준은 운량 데이터를 획득한 기상청 사이트에서 제안하는 방법과 동일하다. Bon 파일 생성 시 그림4와 같이 입력 데이터들을 존재 여부를 확인하도록 하여 혹시나 제외된 데이터가 있다면 확인할 수 있도록 하였다.

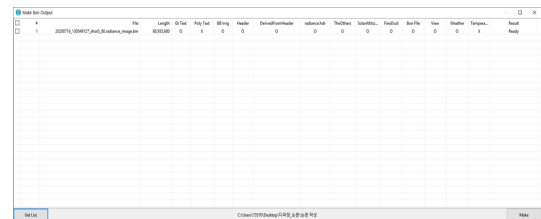
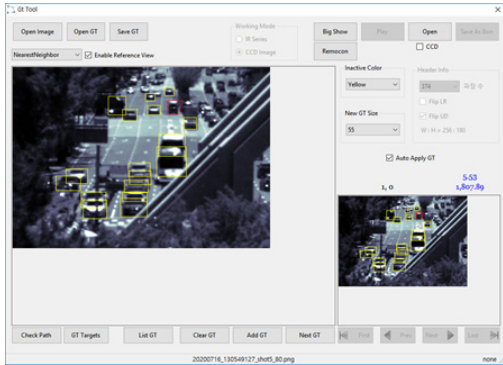


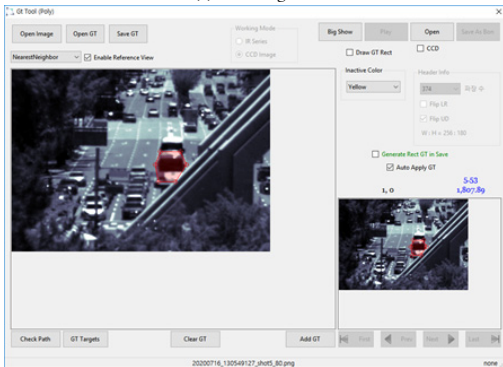
그림 4. Bon 파일 생성 프로그램 화면
 Fig. 4. Bon File Making Program Screen

나. Ground Truth 툴

일반적으로 딥러닝 학습 정보로 사용되는 사각형 Ground Truth와 Polygonal Ground Truth를 획득할 수 있도록 각각의 툴을 개발하였다. Bon 파일, Bon 파일 생성 전 전처리 단계의 초분광 큐브, 일반 형식의 이미지로 저장된 Broadband 이미지 3가지 모두 입력 데이터로 사용 가능하다. Bon 파일의 경우 Bon 파일 내부의 Ground Truth 레이블 값이 변경되고 그 외의 데이터는 txt 파일을 새로 생성하여 Ground Truth 라벨을 저장하고 Bon 파일 생성 때 사용될 수 있도록 하였다. Bon 파일과 전처리 초분광 큐브 데이터를 입력으로 사용할 때는 초분광 큐브의 특성을 고려하여 원하는 대역의 영상을 선택하여 사용하도록 하였다. 이를 통해 Ground Truth 획득 작업 시 물체의 엣지가 가장 잘 구분되는 영상을 선택하여 작업할 수 있다. 또한 작업의 효율을 높이기 위해 Polygonal 라벨 획득 시 획득한 좌표들 가장 바깥에 위치하는 좌표들을 사용해 사각형 라벨을 자동으로 생성하는 기능을 추가하여 작업의 효율성을 개선하였다.



(a) Rectangular



(b) Polygonal

그림 5. Ground Truth 툴 실행 화면
Fig. 5. Ground Truth Tool Program Screen

다. 데이터베이스 관리

데이터베이스 관리는 관리자를 서버와 클라이언트로 나누고 서버에서 생성한 Bon 파일을 데이터베이스에 업로드, 다운로드, 검색, 확인할 수 있도록 하였다. 클라이언트에서는 검색과 확인 기능만 가능하도록 제한하였다.

검색 시에는 그림6과 같이 데이터 종류, 획득 시간, 장소, 촬영 표적과 같은 일반적인 정보들과 대기 온도, 기상 등의 환경정보를 사용하여 조건에 맞는 데이터를 선별하고 환경정보를 확인할 수 있도록 하였다. 이를 통해 학습 목적에 가장 적절한 데이터를 선별할 수 있을 것이라 기대한다.

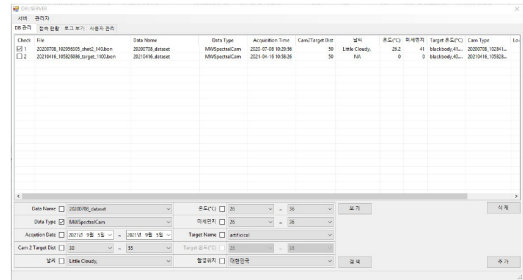


그림 6. 데이터베이스 검색 실행 화면
Fig. 6. Database Search Program Screen

III. 결 론

딥러닝 학습 데이터에 적절한 초분광 영상 데이터를 관리하기 위해 초분광 영상 관리 소프트웨어를 개발하였다. 초분광 영상 데이터의 특징을 고려하여 초분광 큐브에 적절한 Ground Truth 툴을 개발했다. 그리고 가시광 영상과 달리 환경정보에 의해 데이터가 크게 변한다는 특징을 고려하여 환경정보를 함께 관리하는 데이터베이스 관리 기능을 개발하였다. 해당 관리 기능은 환경정보 조건에 따라 검색하는 기능이 있어 딥러닝 학습 목적에 맞는 데이터 생성에 도움을 줄 것이라 기대된다.

추후 초분광 영상을 촬영할 경우 참고용으로 가시광 또는 다른 대역의 적외선 영상을 촬영한다는 점을 고려하여 한 영상에서 획득한 Ground Truth 라벨을 다른 이종 카메라 영상에서 활용할 수 있도록 카메라 좌표계를 변환하는 기능을 추가할 계획이다.

References

- [1] KIM, Yeon-Su. "Military Applications of Hyperspectral Imaging Technique". Optical Science and Technology, Vol. 16, No. 4, pp. 36-4, 2012.
- [2] Eunsoo Park, Byoung-Kwan Cho. "Development of Drought Stress Measurement Method for Red Pepper Leaves using Hyperspectral Short Wave Infrared Imaging Technique", Protected Horticulture and Plant Factory, Vol. 23, No. 1, pp. 50-55, March (2014)
 DOI : <http://dx.doi.org/10.12791/KSBEC.2014.23.1.050>
- [3] Soojin Cho, Byunghyun Kim. "Concrete Surface Degradation Assessment using Hyperspectral Imaging", Magazine of the Korea Concrete Institute, Vol. 29, No. 6, pp. 42-46, 2017.11.
- [4] BAE, Hyungjin, et al. "Development of Non-Destructive Sorting Technique for Viability of Watermelon Seed by Using Hyperspectral Image Processing", Journal of the Korean Society for Nondestructive Testing, Vol. 36, No. 1, pp. 35-44, 2016.
 DOI: 10.7779/JKSNT.2016.36.1.35
- [5] Yun-ji Lim, Sungho Kim "Drug classification using Hyperspectral Image", The Institute of Electronics and Information Engineers, pp. 1005-1008, 2015.
- [6] Beong_Jin Kang, et al "Recent trend of hyperspectral image technology " Institute of Control, Robotics and Systems, pp. 420-421, 2020.7.
- [7] Sungho Kim, Chang Min Ok, "Analysis of Environmental Effects on the Remote Hyperspectral Data Classification" Journal of Institute of Control, Robotics and Systems, Vol. 27, No 8, pp. 602-607, 2021.8.
- [8] Chang, Duhyeuk, Byeonghyeon Jung, and Junyoung Heo. "Classification of Hyperspectral Image Pixel using Optimal Band Selection based on Discrete Range." The Journal of the Institute of Internet, Broadcasting and Communication Vol 21, No 5, pp. 149-154, 2021.
 DOI : <https://doi.org/10.7236/JIIBC.2021.21.5.149>
- [9] Kim, Bal Geum, and Jong Guk Lim. "Discrimination of Internally Browened Apples Utilizing Near-Infrared Non-Destructive Fruit Sorting System." Journal of the Korea Academia-Industrial cooperation Society Vol. 22, No. 1, pp. 208-213, 2021.
 DOI : <https://doi.org/10.5762/KAIS.2021.22.1.208>
- [10] Choi, Kang-In. "Classification of Growth Conditions in Crops Using Hyperspectral Images and Deep Neural Network: Case Study of Paprika Leaf." The Journal of Korean Institute of Information Technology, Vol 17, No. 12, pp. 1-12, 2019.
 DOI : 10.14801/jkiit.2019.17.12.1

저 자 소 개

이 다 빈(정회원)



- 2018년 2월 : 동국대학교 전자전기공학 (공학사)
- 2020년 2월 : 한양대학교 전자컴퓨터통신공학(공학석사)
- 2020년 1월 ~ 현재: LIG넥스원(주) 연구원
- 주 관심분야 : 영상처리, 임베디드 소프트웨어 등

김 흥 락(정회원)



- 1995년 2월 : 대구대학교 전자전기컴퓨터학부 (공학사)
- 1997년 8월 : 대구대학교 전자공학과 (공학석사)
- 1997년 7월 ~ 현재 : LIG넥스원(주) 수석연구원
- 주 관심분야 : 적외선 영상처리, 전원공급기, 영상 처리 시스템 등

박 진 호(정회원)



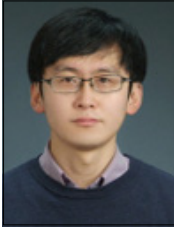
- 2006년 2월 : 부경대학교 전자정보공학부(공학사)
- 2008년 2월 : 부경대학교 정보통신공학(공학석사)
- 2008년 2월 ~ 현재 : LIG넥스원(주) 수석연구원
- 주 관심분야 : 적외선 영상시스템, 디지털 신호처리

황 선 정(정회원)



- 2021년 2월 : 인하대학교 전자공학과 (공학사)
- 2021년 6월 ~ 현재 : LIG넥스원(주) 연구원
- 주 관심분야 : 임베디드 소프트웨어 등

신 정 섭(비회원)



- 2003년 2월 : 충남대학교 컴퓨터과 학 (이학사)
- 2005년 2월 : 충남대학교 컴퓨터과 학 (이학석사)
- 2006년 12월 ~ 2012년 4월 : 국방 과학연구소 선임연구원
- 2013년 5월 ~ 현재 : 국방과학연구소 선임연구원
- 주 관심분야 : 적외선 영상처리 등

※ 본 연구는 국방과학연구소의 지원(UE191116FD)으로 수행된 결과임