

## 딥러닝 모델을 이용한 2D 레고 조립 설명서 생성

안종석\*, 이승현\*, 김철희<sup>o</sup>, 강동희\*

\*동국대학교 AI소프트웨어융합학부,

<sup>o</sup>동국대학교 AI소프트웨어융합학부

e-mail: {jahn, kocan, setg1502, dhk010111}@dgu.edu<sup>o\*</sup>

## Generating 2D LEGO Instruction Manual Using Deep Learning Model

Jongseok Ahn\*, Seunghyeon Lee\*, Cheolhee Kim<sup>o</sup>, Donghee Kang\*

\*Dept. of AI Software Convergence, Dongguk University,

<sup>o</sup>Dept. of AI Software Convergence, Dongguk University

### ● 요약 ●

본 논문에서는 레고(LEGO®) 조립 설명서를 생성하기 위해 딥러닝을 이용한 조립 및 설명서 생성 시스템을 제안한다. 이 시스템은 사용자가 제공한 단일 이미지를 기반으로 레고 조립 설명서를 자동 생성한다. 해당 시스템은 딥러닝 기반 이미지 분할 기술을 활용하여 물체를 배경으로부터 분리하고 이를 통해 조립 설명서를 생성하는 과정을 포함하며, 조립을 위한 알고리즘을 새로 설계하였다. 이 시스템은 기존 레고 제품의 한계를 극복하고, 사용자에게 주어진 부품으로 다양한 모델을 자유롭게 조립할 수 있게 한다. 또한, 복잡한 레고 조립 과정을 간소화하고, 조립의 장벽을 낮추는 데 도움을 준다.

**키워드:** 레고(LEGO®), 조립 설명서(assembly instruction manual), 이미지 분할(image segmentation), 딥러닝 모델(image segmentation), 조립 알고리즘(assembly algorithm)

### I. Introduction

레고(LEGO®)는 모든 연령대에 걸쳐 즐길 수 있는 창의적이고 다양한 완구로 알려졌지만, 시중의 표준 제품과 설명서는 개인의 창의적인 작품을 조립하는 데 한계가 있다. 이러한 문제를 해결하기 위해, 본 연구에서는 사용자가 원하는 단일 이미지를 기반으로 해당 객체의 조립 설명서를 생성하는 2차원 조립 설명서 생성 시스템(2D MAN-GEN)을 제안한다. 이 시스템의 목적은 사용자가 입력한 이미지를 분석하여 최종적으로 레고 조립 설명서를 생성하고 출력할 수 있도록 하는 것이다.

본 논문에서는 물체를 인식하는 과정에 딥러닝 기반 이미지 분할(image segmentation) 기술을 적용하여 조립하고자 하는 물체를 배경으로부터 분리하는 과정을 수행하도록 하였다. 그리고, 적용된 딥러닝 기반 기술을 통해 물체의 윤곽을 검출하는 과정을 정교하게 진행할 수 있으며, 다양한 이미지 상황에도 유연하게 적용할 수 있다는 장점을 적용할 수 있도록 하였다.

이 시스템을 통해 다음과 같은 기존의 레고 제품에 들어있는 설명서의 한계를 극복하고자 한다. 먼저, 기존 레고 제품에 들어있는 조립

설명서는 특정 모델의 조립만을 지원하는데, 이는 사용자에게 주어진 레고 부품을 이용해 다양한 모델을 자유롭게 조립하는 데 제한을 두게 된다. 또한, 사용자가 물체를 표현하는 단일 이미지만을 보고서 레고 모델을 조립하려 할 때, 숙련되지 않은 일반 사용자에게 이를 바탕으로 완성도 높은 작품을 만들기는 어려운 과제이다. 따라서, 본 시스템은 조립 설명서를 제공함으로써 이러한 어려움을 해결하고자 한다.

이 연구의 활용 방안과 기대 효과는 다음과 같이 나열해 볼 수 있다. 첫째, 사용자는 이 시스템을 통해 자신의 창의적인 아이디어를 현실의 물건으로 생성하는 데 도움을 줄 수 있고, 창의적 표현을 확장해 줄 것으로 기대된다. 둘째, 이미지로부터 설명서를 생성하는 기능은 사용자가 조립품을 전시하거나 놀이 도구로 사용하는 등 다양한 요구를 만족시키고, 조립의 장벽을 낮추는 데 도움을 준다. 셋째, 이 모델은 LEGO 조립을 통해 문제해결 능력과 공간 지각 능력을 향상하는 교육적 가치를 가질 것으로 기대된다. 이러한 방식으로, 본 시스템은 LEGO 조립의 새로운 방식을 제시하며, 사용자에게 더욱 풍부하고 만족스러운 경험을 제공할 수 있을 것으로 기대된다.

## II. Preliminaries

### 1. Related works

레고의 조립 설명서와 조립 순서라는 관점에서 수행한 연구가 다양하게 존재하며, 특히 딥러닝을 적용해 레고 조립이라는 복잡한 작업을 수행하는 관련 연구를 찾아보아 진행한 연구와의 연관성을 살펴본다.

**DQN 레고 퍼즐 생성.** 본 논문에서 접근하고 하는 바와 같이 2차원 형태의 레고 조립품을 구현하는 과정에서 강화학습 모델 중 하나인 DQN(Deep Q-Network)을 사용한 연구를 찾을 수 있었다 [1]. 해당 시스템은 레고 조립을 위한 환경을 이미지 그리드와 조립 구성도의 조합으로 설정하고, 레고 부품 선택을 action으로, 완성도를 높이는 정도를 기반으로 한 reward 시스템을 도입했다. 본 논문에서는 DQN이 상대적으로 많은 메모리를 소비하고 학습 속도가 느리다는 점과 사용할 수 있는 action의 수가 제한적이어서 레고 조립의 복잡성을 완전히 해결하기에는 부족할 수 있다는 점을 확인하였고, 이러한 점들을 고려하여 레고 조립 과정을 더욱 효율적으로 최적화하는 알고리즘을 설계할 수 있도록 하였다.

**조립 설명서를 이용한 레고 조립** 본 논문에서는 물체를 나타내는 레고 완성품만이 아닌 그것을 조립하는 과정을 담고 있는 설명서를 최종적으로 도출하는 과정을 연구한다. 이와 관련되어 사람이 읽는 이미지 기반 조립 설명서를 기계가 해석할 수 있는 지침으로 변환하는 연구를 찾아볼 수 있었으며 [2], 해당 모델은 각 단계에서 설명서를 읽고, 현재 형태에 추가될 부품을 찾아내며, 조립될 3D 자세를 추론하게 된다. 해당 시스템과는 다르게 본 논문은 부품들의 위치와 종류, 색상이 순차적으로 표현된 정보를 사람이 읽을 수 있는 형태로 최종적으로 비꾸어 사용자에게 제공함을 목적으로 하고 있으며, 이를 사람이 읽을 수 있는 이미지로 제공함을 목적으로 한다.

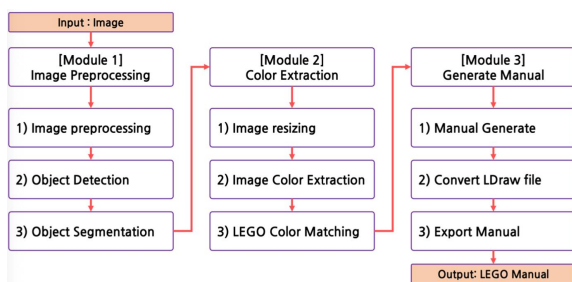


Fig. 1. 2D MAN-GEN System Architecture

## III. The Proposed Scheme

### 1. Methods

이미지를 입력받아 레고 조립 설명서를 출력하는 2D MAN-GEN의 전체 시스템은 Fig. 1에서 볼 수 있듯이 총 세 부분의 모듈로 나누어 볼 수 있으며, 각각의 모듈마다 일련의 작업을 수행하게 된다.

각 모듈의 마지막 단계에서 나온 출력은 다음 단계의 모듈에 입력으로 주어지게 되며, 최종적으로 사람이 읽을 수 있는 이미지 형태의 조립 설명서가 마지막 모듈에서 출력되게 된다. 각 모듈의 세부 동작은 다음과 같이 설명할 수 있다.

#### 1.1 Image Preprocessing

이미지 전처리(image preprocessing) 단계에서는 사용자로부터 입력받은 원본 이미지로부터 원하는 객체를 이미지에서 얻어내는 작업을 수행하게 된다. 원하는 객체를 제외한 배경은 이미지 분할 딥러닝 모델 중 하나인 DeepLabV3+를 이용하여 제거하였으며[3], 사용된 딥러닝 모델은 오픈 소스 시맨틱 분할 툴박스인 ‘MMSegmentation’의 사전 학습된 모델을 사용하였다 [4]. 또한, 원하는 객체의 클래스를 사용자로부터 입력받으며, 입력받은 객체의 클래스 이름은 사전 학습된 딥러닝 모델이 인식할 수 있는 20가지 클래스 범주에 포함된다. 해당 단계의 마지막에서는 배경은 투명한 PNG 형태의 전처리된 객체 이미지를 출력하게 된다.

#### 1.2 Color Extraction

색상 추출(color extraction) 단계에서는 앞서 전처리된 이미지의 크기를 조정된 뒤, 각 픽셀의 색상을 추출하고 가지고 있는 레고 부품의 색상과 대조하여 조립될 레고 색상을 결정하게 된다. 해당 단계의 핵심 기능을 순차적으로 설명하면 다음과 같다. 먼저, 객체 외곽선 검출 기능으로, 조립하고자 하는 객체의 색상만 처리하기 위한 기준으로 외곽선을 검출하는 과정을 수행하게 된다. 다음으로, 배경값 처리 기능으로, 외곽선 외부의 투명한 배경값에 대한 색상 추출 제한한다. 마지막으로, 레고 색상 매칭 알고리즘으로, 추출한 색상 정보(R, G, B) 값과 레고 블록(R, G, B)값과의 차이가 적을, 즉 가장 유사한 색상을 레고 색상 테이블에서 탐색하여 색상 행렬에 저장한다. 이때, 각 RGB 색상 차이를 유클리드 거리로 계산 비교하였다. 해당 단계의 마지막에서는 객체를 조립할 레고 부품의 이미지 색상이 결정된 색상 행렬을 출력하게 된다.

#### 1.3 Generate Manual

설명서 생성(generate manual) 단계에서는 앞서 얻은 색상 행렬을 이용하여 조립 알고리즘을 거쳐 레고 조립 순서를 생성한 뒤, 조립 과정 정보를 활용해 최종적으로 레고 조립 설명서를 제작해 출력한다. 해당 모듈의 조립 알고리즘을 설계하며 발견한 문제점으로 블록을 조립할 때 옆 블록과 결합되는지 확인해야 함을 발견하였다. 옆 블록과 연결되지 않을 경우, 옆 블록과 제대로 결합되지 않았을 때 모형을 두 개 이상의 부분으로 나뉘질 수 있다. 이는 블록이 지그재그로 맞물리지 못하고, 경계가 일직선으로 이어지기에 발생하게 되는데, 경계선이 일직선으로 이어진다는 의미는 블록이 구분되는 지점을 위 또는 아래에서 연결해 주는 블록이 없다는 것이다. 따라서, Fig 2에서 확인할 수 있는 것처럼 모형을 안정적으로 결합되려면 블록들의 경계선이 일직선으로 이어지는 경우가 없어야 한다.

이러한 문제를 해결하기 위해서 블록의 조합을 무작위로 결정하고 검사하는 방법을 택하였다. 이는 같은 블록만을 사용하는 횟수를 줄일 뿐만 아니라, 블록의 경계를 무작위로 분배하여 블록이 좀 더 잘 조립될 수 있도록 하는 의도로 고안했다. 이 방법에서는 이전 단계에서 생성된 색상 행렬의 왼쪽 하단부터 오른쪽 상단 방향으로 순회하는 과정을 거친다. 순회 과정에서 동일한 색상으로 이루어진 연속된 픽셀들을 하나의 그룹으로 취합하고, 그 그룹에서 나올 수 있는 블록의 조합 중 하나를 무작위로 선택한다. 블록의 조합은 조립에 사용될 블록의 크기(1, 2, 4, 6)로 이루어져 있으며, 1로만 이루어져 있는 경우는 발생하지 않는다. 블록의 조합이 선택되었다면 조합 내 블록마다 블록 그룹이 시작된 지점을 기준으로 해서 블록 위치를 계산하고 색상, 크기와 함께 리스트에 저장한다.

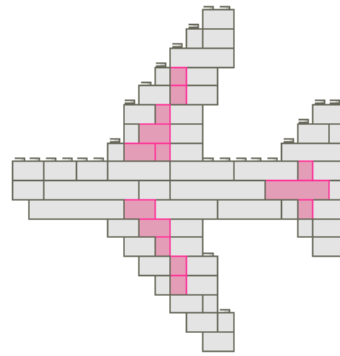


Fig. 3. 비행기 레고 조립 모형

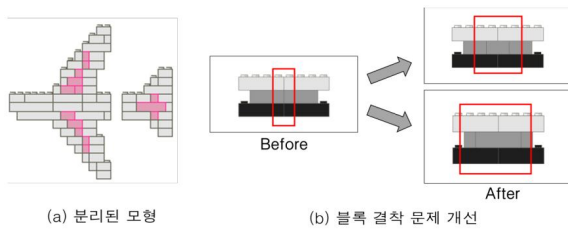


Fig. 2. 블록 결합 아이디어

오른쪽 상단까지 순회가 끝났다면 블록이 제대로 조립되었는지 확인한다. 조립이 잘 되었는지 확인하는 방법은 앞서 말한 대로 블록의 경계선이 일직선으로 일치하는지 확인하는 방법이다. 만약 경계선이 일치하는 부분이 없다면 그대로 리스트를 반환하면 되는 것이고, 일치하는 부분이 존재한다면 앞서 진행된 알고리즘을 다시 진행한다. 반환한 리스트의 경우 사용자의 필요에 따라 JSON, XML 파일로 저장하여 향후 수정하거나, 별도의 파일로 보관할 수 있도록 할 수 있다.

이후, 반환한 리스트나 저장한 JSON, XML 파일을 불러와 사람이 쉽게 이해할 수 있는 이미지 기반 설명서를 출력할 수 있으며, 이때 설명서를 출력하기 위해서는 LDraw 파일 형식에 맞게 내용을 변환할 필요가 있다. LDraw는 컴퓨터에서 3D로 레고 블록 모형을 모델링하기 위한 오픈 소스로 현실에서 조립한 모형 모델을 문서화하거나, 조립 설명서를 만들 수 있는 특징을 가지고 있다. LDraw 파일 형식에 맞게 리스트에 저장된 정보들이 변환되었다면 이를 LDraw 파일로 저장하고, 이 저장된 LDraw 파일을 이용하여 레고 설명서를 출력하기 위한 LDraw 에디터인 LPub3D에서 설명서를 출력한다.

## 2. Results

설계한 2D MAN-GEN을 이용하여 구현한 모델을 통해 생성한 설명서의 모습을 Fig 3에서 확인할 수 있다. 설명서 내 페이지마다 현재 페이지 숫자와 함께 필요한 블록의 종류, 조립했을 때 예상되는 모형의 모습으로 구성되어 있다.

이 모델의 장점으로는 시중에 판매되는 특정 레고 모형 외에도 다양한 모양을 가진 모형을 조립할 수 있게 한다는 점이다. 또한, 사용하지 않는 블록을 재활용할 수 있는 것이다. 단점으로는 사전 학습된 DeepLabV3+ 모델이 20가지 클래스에만 이미지 분할을 진행할 수 있는 점이다. 또한, 블록 조합을 무작위로 선택하는 점으로 인해 같은 이미지로 설명서를 여러 번 출력했을 경우 블록 개수와 설명서 출력 시간이 일관적이지 않다. 이에 따라 입력되는 이미지의 종류가 제한되고, 무작위로 블록 조합을 선택한다는 점 때문에 블록 그룹이 커질수록 조합의 수 또한 커져서 연산량이 늘어나고, 설명서 생성 시간이 오래 걸리게 된다는 문제점을 확인할 수 있었다.

## IV. Conclusions

입력받은 이미지 내 존재하는 객체를 딥러닝 모델을 이용해 인식하고, 설계한 조립 알고리즘을 통해 모형으로 조립할 수 있는 설명서를 출력하는 과정을 진행하였다. 그리고 최종적으로 설명서를 출력하는 모습을 출력하는 모습을 확인할 수 있었다. 이 과정으로 인해 사용자가 원하는 물체를 조립할 수 있음을 확인할 수 있었으며, 해당 시스템이 지는 한계도 살펴볼 수 있었다. 향후 연구로 조립 알고리즘을 개선할 필요가 있으며, GUI 기반의 프로그램 등을 통해 사용자가 쉽게 설명서를 생성할 수 있도록 해당 시스템이 제공되도록 할 계획이다.

## ACKNOWLEDGEMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업 지원을 받아 수행되었음(2023-0-00049).

## REFERENCES

- [1] Park, C., Yang, H., & Min, K. (2020). Generating 2D lego compatible puzzles using reinforcement learning. IEEE

Access, 8, 180394-180410.

- [2] Wang, R., Zhang, Y., Mao, J., Cheng, C. Y., & Wu, J. (2022, October). Translating a visual lego manual to a machine-executable plan. In European Conference on Computer Vision (pp. 677-694). Cham: Springer Nature Switzerland.
- [3] Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European conference on computer vision (ECCV) (pp. 801-818).
- [4] MMSegmentation Contributors. (2020). MMSegmentation: OpenMMLab Semantic Segmentation Toolbox and Benchmark. Retrieved from <https://github.com/open-mmlab/mms Segmentation>