

골프 코스 시뮬레이션을 위한 딥 러닝 기반 이미지 데이터 처리 기법

김승현⁰, 최원제*, 우홍욱*

⁰성균관대학교 소프트웨어학과,

*성균관대학교 소프트웨어학과

e-mail: {kimsh571, wjchoi1995, hwoo}@skku.edu

Deep Learning-based Image Data Processing for Golf Course Simulation

Seunghyun Kim⁰, Wonje Choi*, Honguk Woo*

⁰Dept. of Computer Science and Engineering, Sungkyunkwan University,

*Dept. of Computer Science and Engineering, Sungkyunkwan University

● 요약 ●

본 논문에서는 골프 코스 시뮬레이션을 위해 수집된 데이터의 정제 및 처리에 요구되는 딥 러닝 모델과 모델 적용 과정에 대해서 논의한다. 최근 스크린 골프 시장의 확대와 골프 시뮬레이터 기술의 발전으로 위성 이미지, 항공 촬영 이미지, 공간 정보 시스템 (GIS) 등 다양한 데이터 소스로부터 골프 코스에 대한 정보를 수집에 대한 요구가 증가하였다. 이번 연구에서는 이러한 데이터 소스로부터 생성된 원시 데이터를 최적의 시뮬레이션 입력으로 변환하기 위한 컴퓨터 비전 기법과 딥 러닝 모델 구조에 대해서 검토한다. 특히, 데이터에서 골프 코스 시뮬레이션에 요구되는 메타 데이터를 도출하기 위해 코스 분할(Segmentation)과 코스 오브젝트 분류(Classification) 모델을 적용하는 과정을 다룬다. 이를 통해, 본 연구는 골프 코스 시뮬레이터의 개발 과정에서 중요한 기술 요소를 제공하며, 이는 시뮬레이션의 정확도와 골프 코스의 다양성을 증진시키는데 기여한다.

키워드: 골프 코스 시뮬레이션(Golf Course Simulation), 분할(Segmentation), 분류(Classification)

I. Introduction

스크린 골프, 온라인 게임, 골프 자세 모니터링 등 다양한 형태의 골프 애플리케이션의 활용성이 증가함에 따라, 골프 코스 시뮬레이션 기술도 발전하고 있다. 이러한 코스 시뮬레이션은 골프 훈련 방식을 혁신할 뿐 아니라, 골프 코스 설계자에게는 신규 코스 레이아웃의 기초 정보를 제공한다. 또한, 스크린 골프 환경에서는 고객에게 다양한 골프 코스 환경을 빠르고 효율적으로 제공하는데 맞춤형 기반 기술로 활용 된다.

본 논문은 골프 코스 정보를 위성 이미지, 항공 촬영 이미지, 공간 정보 시스템(GIS) 등 다양한 소스로부터 수집된 이미지 데이터를 딥 러닝 모델과 컴퓨터 비전 기법을 이용해 골프 코스 시뮬레이션에 적합한 형태로 변환하는 방법을 탐구한다. 특히, 골프 코스 시뮬레이션에 필요한 메타 데이터를 추출하기 위해 코스 이미지 분할(Segmentation)과 오브젝트 분류(Classification) 모델의 적용 방식에 대하여 논의한다. 또한, 코스 이미지 분할과 오브젝트 분류 결과가 골프 코스의 특징적인 요소들을 어떻게 도출해낼 수 있는지 검토한다.

본 논문에서 실험된 사전 훈련된(pre-trained) 딥 러닝 기반 비전 모델들은 미세 조정 학습(fine-tuning)을 통해 골프 코스 시뮬레이션의 정확도를 향상시키고, 다양한 골프 코스의 특성을 재현하는 데 도움을 줄 것으로 기대된다. 이를 통해, 본 연구는 기존의 골프 코스 시뮬레이션 개발 접근법에 새로운 관점을 제시하고, 향후 이 분야의 연구와 개발에 중요한 통찰력을 제공하는 데 기여하고자 한다.

II. Proposed Scheme

본 장에서는 딥러닝 사전 모델(pre-trained deep learning model)을 활용하여, 총 600여장의 골프 코스 이미지 데이터로 미세 조정 훈련(fine-tuning)된 분할 및 오브젝트 분류 모델의 정확도에 대해서 논의한다.

1. 골프 코스 분할(Golf Course Segmentation) 모델

골프 코스 분할 데이터 생성 태스크는 골프장 전체 코스 및 단일 홀 코스의 상세한 지형정보를 포함한 맵 생성을 목적으로 한다. 일반적으로 골프장은 18개의 홀로 구성되어 있으며, 각 홀은 페어웨이(Fairway), 그린(Green) 그리고 티(tee)와 같은 의미가 다른 영역을 포함한다. 본 연구에서는 이와 같은 영역의 정보를 포함한 맵 생성 태스크를 기존 컴퓨터 비전 분야에서 주로 다루는 의미론적 분할 문제로 치환한다.

주어진 이미지 I 를 입력으로 받아 각 픽셀 $i \in I$ 에 대하여 해당 픽셀의 골프 코스를 구성하는 영역 레이블 $l \in L$ 를 예측하는 매핑 함수 $f: I \rightarrow L$ 를 찾는 것이다. 데이터셋 $\{(I_1, L_1), (I_2, L_2), \dots, (I_n, L_n)\} \in D$ 에 대하여 실제 맵의 영역 레이블 L_j 와 예측값 $f(I_j)$ 사이의 차이를 최소화하는 함수 f 를 학습하는 손실 함수는 다음과 같이 정의된다.

$$L_{CrossEntropy} = - \sum_{(i, L) \in D} \sum_{i \in I} L(i) \log f(i)$$

여기서 $L(i)$ 는 픽셀 i 의 실제 레이블을 의미한다. 테이블 1과 그림 1은 골프 코스 분할 모델 학습을 위한 데이터셋과 분할 모델의 예시를 나타낸다.

Table 1. 골프 코스 분할 Dataset

| Item | Value |
|-----------|--|
| 영역 레이블 수 | 7개 |
| 레이블 종류 | 배경(background), bunker, 페어웨이(fairway), 그린(green), 그라운드(ground), 호수(lake), 티(tee) |
| 학습 데이터 수 | 211개 |
| 평가 데이터 수 | 23개 |
| 입력 이미지 크기 | 512x512 |

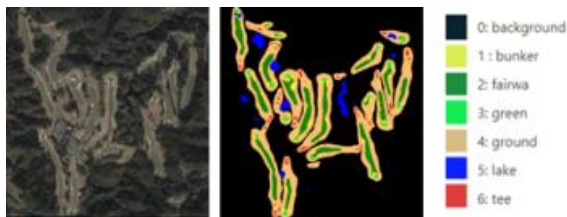


Fig. 1. 골프 코스 분할 데이터 예시 (왼쪽: 입력 이미지, 중앙: 분할 맵, 오른쪽: 분할 레이블)

본 연구의 사전 훈련 모델로는 이미지 분석을 위한 Transformer[6] 기반의 새로운 구조인 Swin-Transformer를 사용한다. 이는 이미지 분류, 의미론적 분할 등의 작업에 있어 매우 효과적이며, 다양한 벤치마크에서 뛰어난 성능을 보여준 것으로 알려져 있다. 비교군으로는 일반 이미지에 대한 의미론적 분할 모델로는 DeepLabV3+[1], HRNetV2+OCR[2] 그리고 Swin-Transformer[3]를 적용한다. 또한, 위성 이미지에 대한 의미론적 분할 모델은 FarSeg[4]과

PointFlow[5]를 적용한다.

Table 2. 사전 모델 미세 조정 성능 비교 표

| 사전 모델 | 성능(mIoU) |
|------------------|----------|
| DeepLabV3+ | 25.4 |
| HRNetV2+OCR | 50.9 |
| Swin-Transformer | 63.98 |
| FarSeg | 59.08 |
| PointFlow | 22.21 |

테이블 2는 다양한 사전 모델을 사용한 미세 조정 학습 한 결과를 나타낸다. 성능 비교에 사용된 평가 지표는 의미론적 분할에서 일반적으로 사용되는 Intersection over Union (IoU)를 사용한다. IoU는 두 개의 영역, 일반적으로 실제 물체의 위치를 나타내는 ground truth와 예측된 결과의 겹치는 영역을 측정하는 방법이다. 가장 높은 성능을 보이는 Swin-Transformer 모델은 'Window based self-attention mechanism'을 도입하여, 이미지를 고정 크기의 non-overlapping patch들로 나눈다. 이러한 작은 window 내에서만 self-attention 연산을 수행함으로써, 계산 복잡도를 줄이고, 동시에 local 정보를 잘 활용한다. 또한, 계층적인 구조를 통해 국소적인 정보를 전체적인 맥락 정보와 통합하도록 하여 local 피쳐 정보뿐만 아니라 이미지 전체의 맥락을 고려할 수 있도록 한다. 따라서 전체 골프장 및 단일 코스를 구성하는 영역들에 대한 이해를 바탕으로 골프 코스 분할 맵 생성 태스크에 최적화되어 높은 성능을 보인다.

2. Golf Course Object Classification(골프 코스 오브젝트 분류) 모델

골프장 전체 코스 및 단일 홀 코스에 포함되는 주변 지형물 정보를 포함한 맵 생성을 목적으로 한다. 일반적으로 골프장 주변 지형물에는 나무, 건물 등과 같은 오브젝트가 존재한다. 본 연구에서는 이와 같은 지형물의 정보를 포함한 맵 생성 태스크를 기존 컴퓨터 비전 분야에서 주로 다루는 오브젝트 분류 문제로 치환한다.

일반적으로 오브젝트 분류 문제의 목적(Objective)는 다음과 같이 표현된다.

$$\arg \min_{\theta} \sum_i^n L(g(I_i; \theta), Y_i), \theta \text{는 오브젝트 분류 모델 } g \text{의}$$

파라미터, I_i 는 학습 데이터셋의 이미지, Y_i 는 이미지에 대한 ground truth로 g 의 예측값과 Y 간의 차이를 최소화하는 방향으로 θ 한다. 이때 손실 함수 L 은 일반적으로 분류 오차를 계산하는 Cross Entropy 손실 함수와 Bounding Box의 위치 오차를 계산하는 Regression 손실 함수로 구성된다. 테이블 3과 그림 2은 골프 코스 오브젝트 분류 모델 학습을 위한 데이터셋과 분할 모델의 예시를 나타낸다.

Table 3. 골프 코스 오브젝트 분류 Dataset

| Item | Value |
|-----------|---|
| 영역 레이블 수 | 11개 |
| 오브젝트 종류 | Green, Bunker, Tree, Forest, Building, Car, Boat, Lake, Football Park, American Football, Warehouse |
| 학습 데이터 수 | 326개 |
| 평가 데이터 수 | 10개 |
| 입력 이미지 크기 | 224x224 |

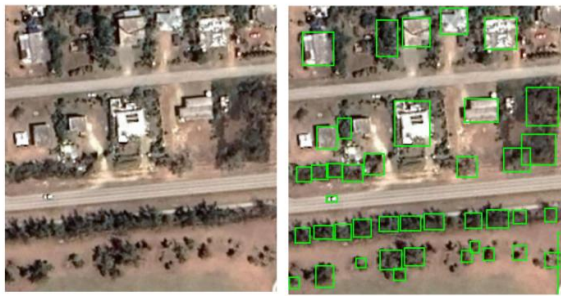


Fig. 2. 골프 코스 오브젝트 분류 데이터 예시 (왼쪽: 입력 이미지, 오른쪽: 오브젝트 분류 맵)

본 연구의 사전 훈련된 모델로는 ResNet[7]기반의 Faster R-CNN[8] 오브젝트 분류 모델을 사용한다. Faster R-CNN 모델은 ResNet을 backbone으로 사용하며, 이미지에서 다양한 크기와 형태의 오브젝트를 효과적으로 검출할 수 있다. 사전 학습 모델은 COCO(Common objects in context)[9] 데이터셋을 통해 학습된 모델을 사용한다.

Table 4. 사전 모델 미세 조정 후 각 오브젝트에 대한 성능 비교 표

| 오브젝트 | 미세조정 성능(AP) |
|-------------------|-------------|
| Green | 100.00 |
| Bunker | 100.00 |
| Tree | 53.11 |
| Forest | 68.65 |
| Building | 70.96 |
| Car | 73.56 |
| Boat | 51.38 |
| Lake | 100.0 |
| Football Park | 66.67 |
| American Football | 50.50 |
| Warehouse | 37.50 |
| Total | 70.21 |

테이블 4는 사전 학습 모델을 미세조정 했을 때, 각 오브젝트에 대한 성능을 비교한 표이다. 성능 지표는 mAP(mean Average Precision)를 사용한다. Average Precision (AP)는 정밀도-재현율 곡선 (Precision-Recall curve) 아래의 영역을 나타내는 메트릭으로, 특정 클래스에 대한 모델의 성능을 나타 측정하는 데 사용된다. 이때, 정밀도 (Precision)는 모델이 True Positive로 예측한 샘플 중 실제로

True Positive인 샘플의 비율이며, 재현율 (Recall)은 실제 True Positive 샘플 중 모델이 True Positive로 예측한 샘플의 비율을 의미한다. 전체 오브젝트에 대한 AP 평균은 70.21로 상용화 가능한 수준의 높은 정확성을 보인다.

III. Conclusion

본 논문에서는 골프 코스 이미지 데이터를 사전 훈련된 최신 딥러닝 모델을 미세 조정하여 골프 코스 시물레이션 요구되는 메타 데이터 생성 기술 구조에 대해서 논의하였다. 실험된 코스 이미지 분할(Segmentation)과 오브젝트 분류(Classification) 모델은 스크린 골프 환경에서 사용자 맞춤형 코스 생성과 같이 새로운 골프 코스 시물레이션 기술 발전에 중요한 역할을 할 것으로 기대된다.

ACKNOWLEDGEMENT

본 연구는 2023년도 중소벤처기업부의 기술개발혁신사업 지원에 의한 연구임 (S3151747). 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW컴퓨팅산업원천기술개발사업의 연구결과로 수행되었음(2021-0-00875).

REFERENCES

- [1] 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).
- [2] Sun, K., Zhao, Y., Jiang, B., Cheng, T., Xiao, B., Liu, D., Mu, Y., Wang, X., Liu, W., & Wang, J. (2019). High-resolution representations for labeling pixels and regions. arXiv preprint arXiv:1904.04514.
- [3] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).
- [4] Zheng, Z., Zhong, Y., Wang, J., & Ma, A. (2020). Foreground-aware relation network for geospatial object segmentation in high spatial resolution remote sensing imagery. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4096-4105).
- [5] Li, X., He, H., Li, X., Li, D., Cheng, G., Shi, J., Weng,

- L., Tong, Y & Lin, Z. (2021). Pointflow: Flowing semantics through points for aerial image segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4217-4226).
- [6] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., A. N., Gomez, Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [7] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- [8] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.
- [9] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13* (pp. 740-755). Springer International Publishing.