

## ESL의 YOLOv5: 참여 학습을 위한 객체 감지

파딜라 존에드워드<sup>o</sup>, 이강희(교신저자)\*

<sup>o</sup>숭실대학원 미디어학과,

\*숭실대학원 미디어학과

e-mail: jedpadilla@soongsil.ac.kr<sup>o</sup>, kanghee.lee@ssu.ac.kr\*

## YOLOv5 in ESL: Object Detection for Engaging Learning

John Edward Padilla<sup>o</sup>, Kang-Hee Lee(Corresponding Author)\*

<sup>o</sup>Dept. of Digital Media, Soongsil University,

\*Dept. of Digital Media, Soongsil University

### ● 요약 ●

In order to improve and promote immersive learning experiences for English as a Second Language (ESL) students, the deployment of a YOLOv5 model for object identification in videos is proposed. The procedure includes collecting annotated datasets, preparing the data, and then fine-tuning a model using the YOLOv5 framework. The study's major objective is to integrate a well-trained model into ESL instruction in order to analyze the effectiveness of AI application in the field.

**키워드:** English as a Second Language, ESL, object detection, sight words, YOLOv5

### I. Introduction

There is an increasing interest in experimenting with new techniques to improve English as Second Language (ESL) training and create immersive language learning experiences [1].

The YOLOv5 (You Only Look Once) deep learning model for real-time object identification capabilities, is accurate in identifying and localizing objects within pictures [2]. Visual aids and interactive learning enhance ESL learners' vocabulary, particularly in sight word learning, which focuses on widely used terms [3].

### II. Methodology

The following incorporate the YOLOv5 algorithm for image identification in videos:

1. Data Gathering
2. Data-preprocessing
3. Model Training
4. Integration into ESL

Use YOLOv5 to recognize and highlight sight words in videos, giving ESL learners with visual aid and interactive learning.

### III. Results

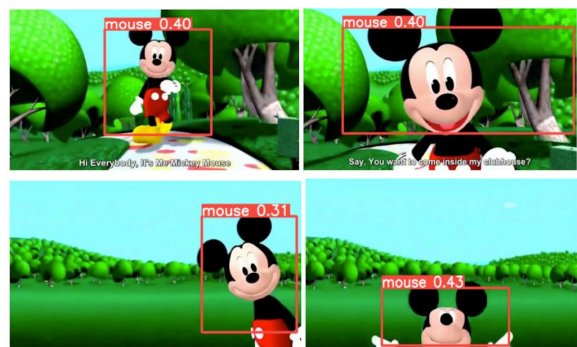


Fig. 1. Screenshots from the video, showing bounding boxes with labels

The YOLOv5 system recognized Mickey Mouse in a video using a dataset of 20 pictures (15 for training and 5 for validation). By extracting bounding box annotation around the subject,

makesense.ai assisted in tagging the photographs. Ultralytics YOLOv5 version 7.0 was used for training, with particular settings such as a batch size of 1 and training for 240 epochs.

Training the YOLOv5 model increases bounding box predictions and object identification accuracy, with positive trends in instance count, accuracy, recall, and mAP metrics. The trained YOLOv5 model correctly spotted and localized occurrences of Mickey Mouse (fig. 1).

#### IV. Discussion

The success of YOLOv5 to implement the technology to teach sight words in ESL can have several advantages and implications.

##### 1. Contextual understanding

Teaching sight words primarily through books might restrict students' understanding of their practical use in real world [4]. YOLOv5 can aid in understanding how words are encountered in real-world circumstances.

##### 2. Vocabulary gain

By associating sight words with visual cues, students can strengthen their understanding and recollection of these keywords while also expanding their vocabulary range [5].

#### V. Conclusion

The effective object identification of the trained YOLOv5 model in the movie highlights its potential to improve ESL instruction by improving vocabulary acquisition, contextual comprehension, and gamification of the learning process for non-native English speakers.

#### REFERENCES

- [1] Angouti, S., & Jahangiri, K. Transforming ESL Teaching Modalities Using Technological Tools. *ELLS*, 1(9), 134, 2019.
- [2] GitHub – ultralytics/yolov5, <https://github.com/ultralytics/yolov5>
- [3] Pulido, D. Modeling the Role of Second Language Proficiency and Topic Familiarity in Second Language Incidental Vocabulary Acquisition Through Reading. *Journal of Research in Language Acquisition*, 36(3), 321-345, 2003.
- [4] Browder, D. M., & Xin, Y. P. A Meta-Analysis and Review of Sight Word Research and Its Implications for Teaching Functional Reading to Individuals with Moderate and Severe Disabilities. *The Journal of Special Education*, 32(3), 1998,
- [5] Tahir, M. H. M., Albakri, I. S. M. A., Adnan, A. H. M., Shaq, M. S. Y., & Shah, D. S. M. The Application of Visual Vocabulary for ESL Students' Vocabulary Learning, 2020.

#### ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2018-0-00209)