Journal of Open Humanities Data

Automatic Detection of Four-Panel Cartoon in Large-Scale Korean Digitized Newspapers using Deep Learning

RESEARCH PAPER

SEOJOON LEE
BYUNGJUN KIM
BONG GWAN JUN
*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

In the realm of cultural and historical studies, the collection of image-based content from big data is a fundamental aspect of data analysis. However, this process is as intricate as extracting resources from vast terrains. Echoing this sentiment, there is a growing appreciation in scholarly circles for "Four-panel Cartoon" (FPC) as a valuable image content source in big data digital newspapers in the Republic of Korea. Yet, identifying these FPCs amidst the vastness of big data archives is an arduous journey, especially given their unstructured image data format — a task both time-intensive and costly. To address this issue, this research paper presents a novel computational FPC detection mechanism: the development of the YOLOv5_FPC model, via fine-tuning the You Only Look Once Version 5 (YOLOv5) deep learning model, tailored precisely for FPC image detection. We applied our YOLOv5_FPC model to the Chosun Ilbo News Library archive (1920–1940) for automatic FPC data mining, spanning 47,777 JPG image files. We identified 1040 FPC objects within 1035 files, which include previously undiscovered FPCs by previous researchers. We provide a detailed description of our methodology, which includes the collection, labeling, training, detection, and distribution of the data we discovered from big data newspaper archives. Our findings, now available as an open-access dataset in the Journal of Open Humanities Data (JOHD) Dataverse, invite discussions among humanities researchers focusing on the culture and history of Korea between 1920 and 1940.

CORRESPONDING AUTHOR:
Bong Gwan Jun
School of Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology, Daejeon, the Republic of Korea
junbg@kaist.ac.kr

KEYWORDS:
big data; object detection; data strategy; four-panel cartoon; digital newspaper; data science

TO CITE THIS ARTICLE:
Lee, S., Kim, B., & Jun, B. G. (2024). Automatic Detection of Four-Panel Cartoon in Large-Scale Korean Digitized Newspapers using Deep Learning. *Journal of Open Humanities Data*, 10: 36, pp. 1–15. DOI: https://doi.org/10.5334/johd.205

1 CONTEXT AND MOTIVATION

In the era of digital transformation, rapid advancements in science and technology influence various spheres of human activity, raising questions about their instrumental values for humanity. As technology advances, digital humanities researchers play an important role in this era across various intersections, including the domains of humanities and Artificial Intelligence (A.I.). This is driven by the public's desire to harness A.I. to fulfill aspirations and address needs. In this context, digital humanities experts, with an understanding of both humanities and A.I., are vital. Their competence is important for curating data, particularly in humanities research, where not all data is suitable for machine learning aimed at humanistic goals in the real world. When data synergizes with algorithms and computational resources, it can amplify its potential. Understanding the processes of data collection, algorithms, curation, analysis, and distribution of their datasets ensures the data's continued relevance and value for humanities purposes. Furthermore, sharing their datasets in Internet databases fosters global interdisciplinary research from diverse backgrounds. In this research paper, we provide a comprehensive account of our data strategy: collecting, labeling, training, employing You Only Look Once (YOLO) algorithm for data mining from big data, and disseminating the dataset we discovered from the Chosun Ilbo newspaper as an Excel file. Also, we demonstrate how open-access digital platforms can foster global research by developing scripts for automatic object detection, with the potential for our computational methodology to be expanded in future research to discovering other types of image sources in big data.

The Chosun Ilbo, a big data newspaper digital archive, commenced publication in 1920 and ceased operations on August 10, 1940, under pressure from the Governor-General of Chosun (Kim, 2005), offers valuable insights for the analysis of culture and history. A key element in these archives is the Four-Panel Cartoon (FPC) in Korea, a common feature in historical newspapers and magazines (Lee, 2005). These FPCs, particularly significant during the Japanese colonial era (1910–1945), are more than just cartoons; they are satirical commentaries, as seen in publications like Chosun Ilbo (*ChosunIlboNewsLibrary*, 2024). FPCs, or "Naecutmanwha" in Korean, offer a window into the culture and politics of the era (Park, 2010). Also, their concise yet effective narrative development and sequential structure have been serialized in various newspapers for decades, capturing significant public attention. These newspaper FPCs hold not only artistic value in the realm of art history but also possess historical values, as their serialization gained momentum during the Japanese colonial period in Korea. An example is the "Meongteongguri" series, which goes beyond mere entertainment to reflect the complexities of life in colonial-era Seoul (Chung, 2016). These periods saw restricted expression, making satire an important form of commentary (Sa, 2009). However, identifying FPCs in vast digital archives spanning decades poses a challenge due to the sheer volume of digital files and the unstructured nature of the image objects in the big data archives. Above all, FPCs manifest in various sizes and shapes, scattered irregularly. Therefore, the development of an efficient object detection algorithm is imperative. To achieve our research goals by addressing this issue, our research employs the YOLOv5 deep learning model, specifically adapted for FPC detection. This model, known as YOLOv5_FPC, identified 1040 FPC images from 47,777 files in the Chosun Ilbo. The authors created an Excel file containing the URLs of the YOLOv5_FPC-detected FPC images, along with publication dates, which is now available on the JOHD Dataverse (Lee et al., 2024a). Any researchers can investigate the FPCs by clicking once on the corresponding URLs in the Excel file, encompassing previously undiscovered FPCs by previous researchers. The data we found encourages novel discourse among humanities scholars interested in Korean culture and history from 1920 to 1940. Our dataset sheds light on how society changed through FPCs. Additionally, by developing scripts for automatic FPC detection using our YOLOv5_FPC weights on the Google Colab platform, this tool simplifies FPC detection for the public, enabling them to detect FPCs in their local computer files. Our interdisciplinary methodology advances culture studies, history studies, digital humanities, and deep learning, enriching academic discourse and achieving humanities objectives.

2 DATASET DESCRIPTION

Repository location https://doi.org/10.7910/DVN/DFVZWE

Repository name Journal of Open Humanities Data Dataverse

Object name "Metadata for YOLOv5_FPC-Detected Images"

Format names and versions Excel and CSV files (each file contains metadata for the YOLOv5_FPC-detected four-panel cartoons, including the respective URLs for the 1035 image files which contain 1040 FPC objects).

Creation dates 2023-02-28–2024-04-15.

Dataset creators Seojoon Lee (training data collection, YOLOv5 fine-tuning, big data collection, data mining, database curation, data analysis, and data conceptualization).

Byungjun Kim (big data collection, database curation, data conceptualization, and data advice).

Bong Gwan Jun (project administration, data conceptualization, and data advice).

Language The Excel and CSV files: Written in English. Discovered FPC images: Mostly Korean, partially Japanese four-panel cartoons.

License CC BY-NC-SA 4.0 DEED (Attribution-NonCommercial-ShareAlike 4.0 International).

Publication date 2024-04-15

3 METHOD

3.1 RELATED WORKS

To select an appropriate deep learning model for our research, we commenced by conducting a comprehensive literature review. Our objective during this process was to identify a model that meets two essential criteria to accomplish our research goal: high accuracy in object detection and expedited detection speed. Traditionally, object detection has been an essential subdomain within computer vision, significantly contributing to a wide range of academic and commercial use cases. Deep learning techniques have proven highly effective for image categorization and object detection tasks within multimedia visuals, consistently yielding significant results on established benchmark datasets (Girshick et al., 2013; Krizhevsky et al., 2017; Redmon et al., 2016). Examples of such applications include face detection (Taigman et al., 2014), medical image detection (Litjens et al., 2017), vehicle detection (Chen et al., 2013), and logo detection(Hoi et al., 2015). We also have observed that there were endeavors that employ object detection in digital humanities. Siddiqui (2024) employed the Google Vision API to identify objects within 105,000 scenes from 15 films directed by Alfred Hitchcock. This approach facilitates the study of films and digital culture by analyzing movie scenes and their representation in online culture. Fiala et al. (2022) developed an algorithm that automatically converts sixteenth and seventeenth-century mensural sheet music into modern notation software formats, achieving a 99% recognition rate in translating musical material into MusicXML format using a four-step optical music recognition approach. Smirnov suggested utilizing deep learning and neural networks to perform automatic object detection on digital images of fine-art paintings. To achieve this, Smirnov employed a Convolutional Neural Network (CNN) that employed the VGG-19 architecture and was optimized using Bayesian techniques and transfer learning. The results of their approach showed that automatic object detection had an accuracy rate of 64% (Smirnov and Eguizabal, 2018). Hodel et al. (2021) researched developing advanced models for handwritten text recognition, enabling the accurate interpretation of historical documents across a variety of handwriting styles from the German Kurrent script spanning over a century. It benefits the public by providing access to historical texts, thus unlocking new possibilities for research and preservation within digital humanities and archival fields.

In recent years, the adoption of the You Look Only Once (YOLO) object detection algorithm has seen rapid growth across diverse academic and professional domains (Cao et al., 2020). Rajeshwari contends that YOLO provides exceptionally high levels of accuracy and speed, surpassing deep learning models such as R-CNN by nearly 1000 times in speed and outperforming the Fast R-CNN model by a factor of 100 in terms of speed. YOLO utilizes a fully convolutional neural network to generate grids of size S*S across an image, along with bounding boxes and associated class probabilities for each grid (Rajeshwari et al., 2019). After comparing pre-existing deep learning models in the literature review, we decided to employ the You Only Look Once Version 5 (YOLOv5) model (Jocher, 2020), given its ability for multi-class object detection with high accuracy and fast detection speed, fulfilling the two fundamental qualifications required for our research.

3.2 METHODOLOGICAL APPROACH

In this methodology section, we outline detailed procedures spanning phases 1 through 9 as presented in Table 1. These phases include training data collection, labeling, fine-tuning, evaluating the YOLOv5_FPC model, big data collection, data mining of the FPCs, database curation, and developing the YOLOv5_FPC-Detector scripts utilizing the Google Colab platform for wider applications.

Table 1 Research phases of this paper.

NUMBER	RESEARCH PHASE
1	Training data collection of Four-Panel Cartoons (FPCs).
2	Labelling process of the training dataset.
3	Fine-tuning of the YOLOv5 Model.
4	YOLOv5_FPC model evaluation: F1-score.
5	Image file collection from the Chosun Ilbo News Library (1920–1940), totaling 47,777 JPG files.
6	Data mining: Deployment of the YOLOv5_FPC model to the 47,777 JPG files from the Chosun Ilbo News Library, to detect FPC image ojects.
7	Database curation: Uploading the Excel and CSV file which contains metadata for URLs of YOLOv5_FPC-detected 1035 images files (1040 FPC objects), which includes preciously undiscovered FPCs, to the JOHD Dataverse (Lee et al., 2024a).
8	Data analysis of the detected FPC objects.
9	Development of the YOLOv5_FPC-Detector script, leveraging the Google Colab platform for enhanced computational efficiency and wider application for the public.

Detecting objects within images is an essential component of computer vision, commonly referred to as object detection. We implemented a pioneering technique that leverages the You Only Look Once (YOLO) algorithm for identifying FPCs in large-scale digital documents. Specifically, we utilized the You Only Look Once Version 5 (YOLOv5) model (Jocher, 2020), a neural network-driven regression approach that facilitates single-stage object detection (Xiao et al., 2020) with high accuracy and rapid detection. YOLO employs a fully Convolutional Neural Network to produce S*S grids across an image, bounding boxes for each grid, and probabilities for each class associated with the bounding boxes (Rajeshwari et al., 2019). It is an object detection algorithm capable of simultaneous detection and classification of objects with high accuracy and rapid speed. We needed to select an agile model that satisfies two critical criteria for our research: high accuracy and fast speed. Thus, after conducting a comprehensive literature review, we determined that this model qualifies for the object detection in our study, which demands high accuracy and rapid detection speed when detecting FPCs in large-scale digital documents. Moreover, YOLO demonstrates proficiency in detecting small objects and exhibits confidence in detecting occluded objects (Rajeshwari et al., 2019), considering the possibility that there could be a presence of small or occluded FPC objects in the scanned and digitized newspapers.

As Figure 1 demonstrates, the initial YOLOv5 model could not detect the FPC image object in the testing dataset. We attributed this to the fact that the initial YOLOv5 model was not pre-trained with FPC image data; the initial YOLOv5 model was trained solely on the Microsoft COCO dataset, which includes only 80 object categories such as cars, bicycles, and wine glasses. To overcome this limitation, we utilized the fine-tuning method from the Ultralytics GitHub website (Jocher, 2020) to train the YOLOv5 model with our training dataset to detect the FPCs in digital documents. To train the model, we created bounding boxes for each shape within the 161 pages of FPC images using the DarkLabel 2.4 software (Darkpgmr, 2021). These bounding boxes were then converted into the YOLOv5 format, along with the corresponding coordinates. We used the Roboflow software (Dwyer et al., 2022) to randomly divide the images into training, validation, and testing sets. The training dataset served to optimize the YOLOv5 model for precise detection of FPCs, while the validation dataset was used for model development and the hyperparameter fine-tuning process during the training. The testing dataset was employed to evaluate the model's overall performance and to mitigate the risk of over-fitting. We divided the 161 images into 113, 24, and 24 files for the training, validation, and testing sets, respectively,

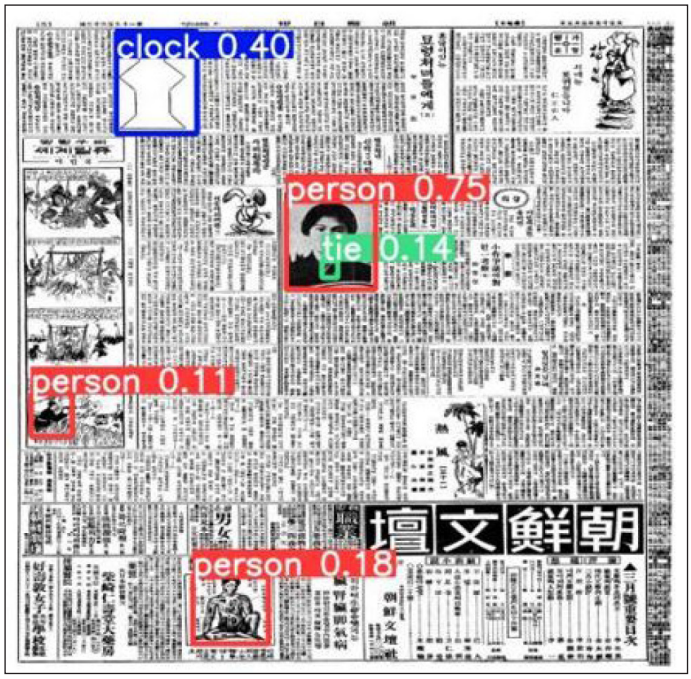


Figure 1 The initial YOLOv5 Model could not detect an FPC.

and saved them in separate directories. In our study, we employed a relatively small number of training, validation, and testing datasets due to the inherent challenge of pinpointing the exact locations of FPCs within the vast expanse of Internet. Our aim was to investigate the feasibility of utilizing a limited amount of training data for the detection of FPC objects. Additionally, we created a label directory that contained the coordinate axis of every bounding box in a text file for each figure, which was used for YOLOv5 training.

Creating high-quality datasets is an important aspect of training and testing deep learning models, and is especially crucial when dealing with complex objects such as FPCs. In this paper, we introduce the "Four-panel Cartoon Image Dataset" (Lee et al., 2023b) — a mix of grayscale and color images from 1920 to 2022, which we used for fine-tuning. The composition includes two categories of images: a 4 × 1 matrix and a 2 × 2 matrix. A 4 × 1 matrix FPC denotes a graphical illustration arranged in a rectangular configuration consisting of four rows and a single column, whereas a 2 × 2 matrix cartoon represents an illustration with two rows and two columns. These configurations offer distinct layouts for visually presenting information, facilitating varied academic and artistic applications. Our decision to exclude the 1 × 4 matrix FPC images was based on the fact that such FPCs are not commonly utilized in Korean cartoons, leading us to conclude that their inclusion in our study would not provide a representative sample of the prevalent usage of FPCs. Our dataset comprises 161 JPG format image files, randomly divided into the training (113 files), validation (24 files), and testing (24 files) set using Roboflow software (Dwyer et al., 2022). Table 2 depicts the division of the dataset into training, validation, and testing sets, depending on the matrix, and the era.

Table 2 Matrix and era of the "Four-panel Cartoon Image Dataset".

SET	FPC MATRIX	COLONIAL ERA	POST-COLONIAL ERA
Training	4 × 1	31	37
	2 × 2	26	28
Validation	4 × 1	12	3
	2 × 2	4	7
Testing	4 × 1	8	6
	2 × 2	6	7

3.3 FINE-TUNING FOR FPC DETECTION OPTIMIZATION

We fine-tuned the initial YOLOv5 model by using the training dataset we collected to optimize the FPC object detection.

In Figure 2, the displayed scores of box_loss, object_loss (obj_loss), and class_loss (cls_loss) provide insights into the model's performance during the training and validation processes. In Arie's explanation, the YOLO loss function comprises three components: box_loss, object_loss, and class_loss. Box_loss represents the loss incurred in bounding box regression, calculated using Mean Squared Error. Obj_loss quantifies the confidence of object presence, referred to as the objectness loss. Cls_loss accounts for the classification loss, computed using Cross Entropy. Arie also explains precision and recall. Precision assesses the model's correctness in making positive predictions, while recall evaluates the model's capacity to accurately detect positive occurrences; these metrics are in a trade-off relationship. The metrics "mAP_0.5" refers to the mean Average Precision (mAP) at an IoU threshold of 0.5, while "mAP_0.5:0.95" denotes the average mAP across various IoU thresholds, ranging from 0.5 to 0.95 (Arie, 2022). In Figure 2, several key points warrant attention. In the train/box_loss and val/box_loss graphs, as the number of epochs increases from 0 to 200 on the x-axis, there appears to be a decrease in the box loss on the y-axis. This decline signifies an improvement in accuracy when predicting bounding box coordinates. In the train/obj_loss and val/obj_loss graphs, as the number of epochs increases from 0 to 200, there is a decrease in the loss values which indicates an improvement of accuracy in the detection of objects. Regarding the train/cls_loss and val/cls_loss graphs, an increase in the epoch count in the x-axis results in a decrease in classification loss of the y-axis. The decrease demonstrates an enhancement in accuracy for classifying objects. In the metrics/precision graph, as the epoch count increases from 0 to 200, the precision score tends to increase as well. In each of the ten graphs depicted in Figure 2, the blue lines labeled as "results" represent performance metrics and outcomes, while the orange dots labeled as "smooth" indicate a smoothed version of the blue curves, helping to reduce noise and fluctuation for easier interpretation.

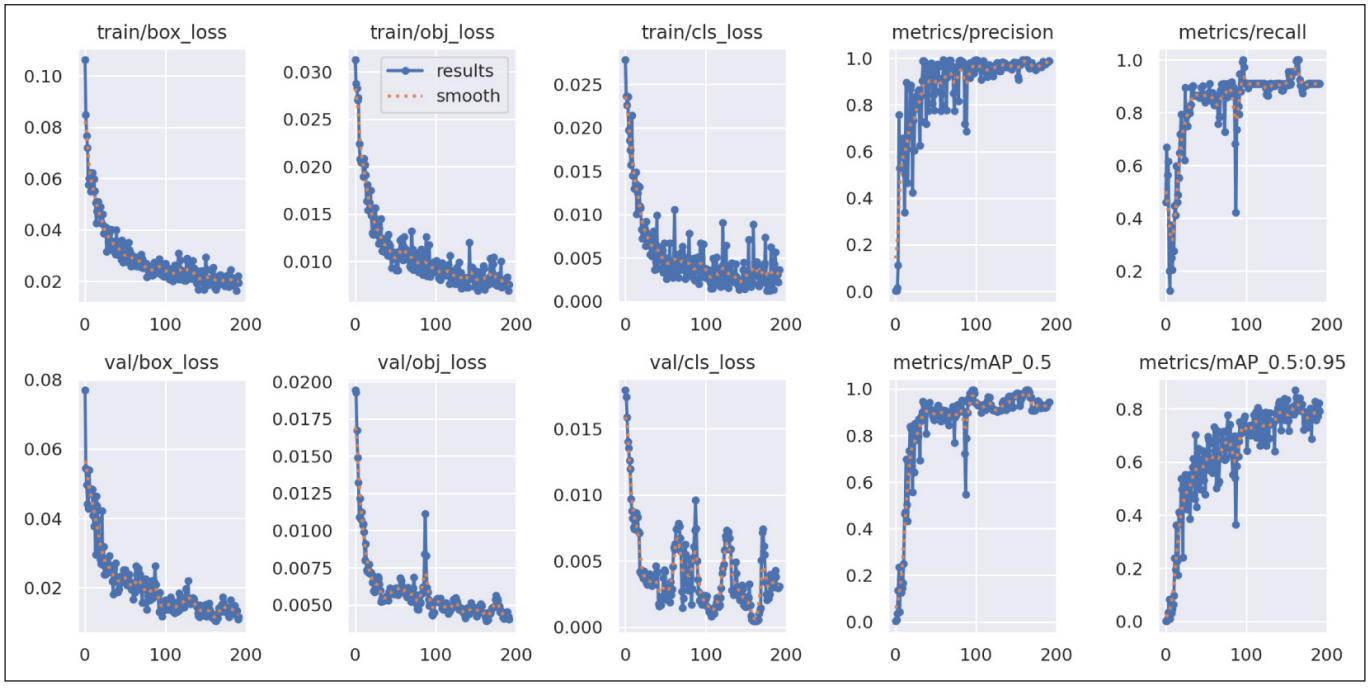


Figure 2 Model performance while fine-tuning.

$$\textbf{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\textbf{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

The F1-score, which unifies precision and recall, serves as the harmonic average of these two evaluation measures. The F1-score takes into account both precision and recall, which are important metrics for evaluating object detection models. Our fine-tuned model, depicted in Figure 3, achieved a high F1-Confidence score of 0.97 for all classes at a confidence threshold of 0.708, signifying accurate and balanced performance in precision and recall across both 4 × 1

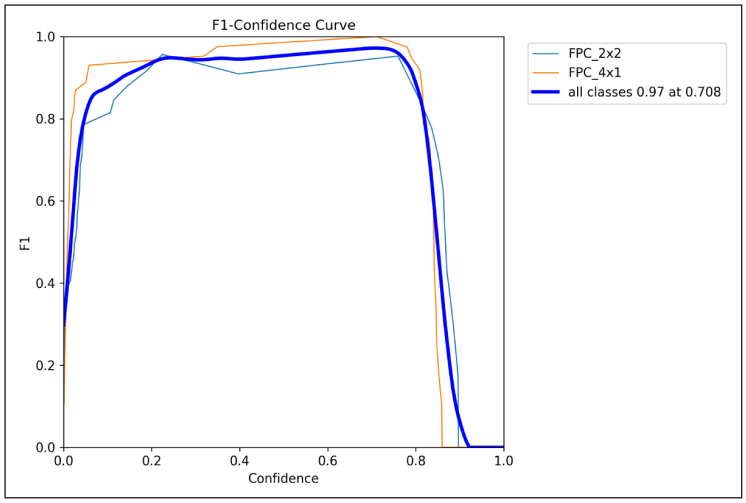


Figure 3 F1-score of our YOLOv5_FPC model.

and 2 × 2 FPC categories. This YOLOv5_FPC model weights will be utilized for data mining on big data newspapers in this research.

3.4 DATA MINING: FPC OBJECT DETECTION IN THE CHOSUN ILBO NEWS LIBRARY

The Excel file depicted in Figure 4 contains columns including "image", "source_xml_file", "publication_day", "id", "type", and "body", representing the metadata of the 47,777 JPG files from 1920 to 1940 in Chosun Ilbo. Table 3 delineates the definitions of the metadata included in this Excel file. The "image" column, which contains the URLs of the scanned and digitized image files from the Chosun Ilbo newspaper database, was utilized to collect 47,777 JPG image files from Chosun Ilbo (*ChosunIlboNewsLibrary*, 2024). In total, we collected 47,777 JPG files from the database (Figure 5), using automation. Leveraging the deep learning model YOLOv5_FPC, FPCs were identified within the image files. The original YOLOv5 model, trained on a suite of 80 conventional objects utilizing the Microsoft COCO dataset, was initially incapable of detecting unfamiliar objects including, but not limited to, FPCs. However, after fine-tuning and retraining using FPC train data gathered by the authors in this research, the modified YOLOv5_FPC model adeptly identified FPCs. Utilizing the Chosun Ilbo dataset, our advanced deep learning model detected a total of 4508 image files. Of these, 1035 files encompassed 1040 FPC objects in the 47,777 image files, which underwent manual verification by the authors to confirm their authenticity as FPCs. Noteworthily, the YOLOv5_FPC model exhibited prowess in detecting multi-class FPCs even when both 4 × 1 and 2 × 2 matrix structures coexisted within a singular file. In the 4508 detected image files, along with accurately identified FPC objects, there were also false positive redundant objects, encompassing instances of the vertical Chosun Ilbo title — written in Chinese characters, and vertical advertisements. In the realm of deep learning object detection, a "false positive" is an instance wherein the model incorrectly

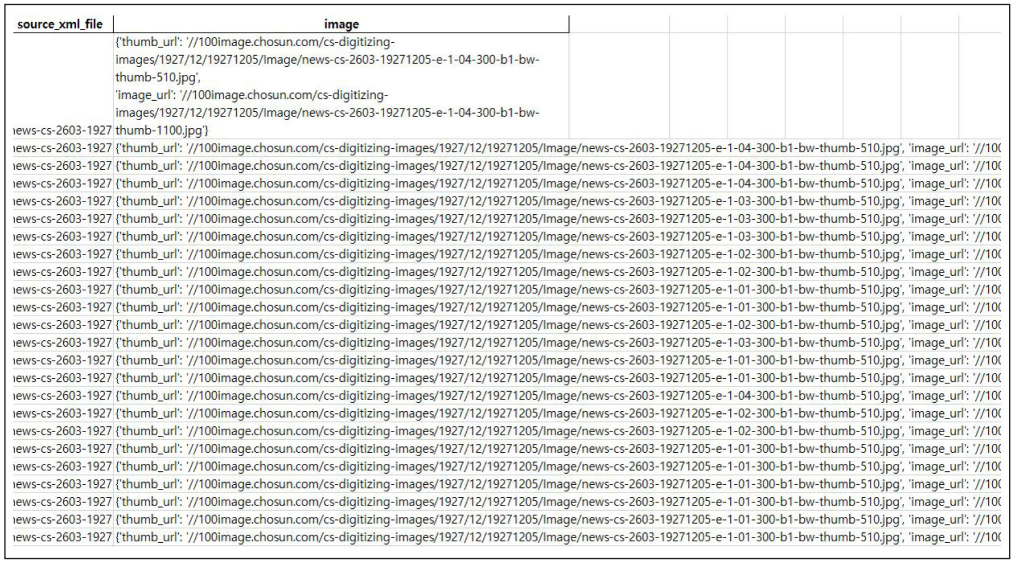


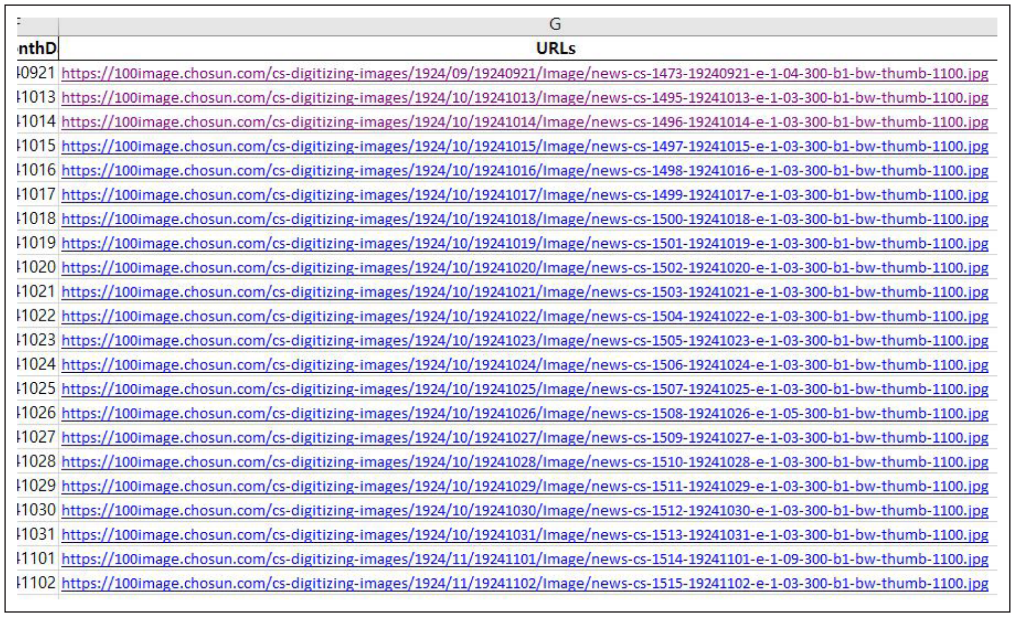
Figure 4 Chosun Ilbo News Library newspaper metadata (1920–1940) (*ChosunIlboNewsLibrary*, 2024).

COLUMN NAMES	DEFINITION
id	The unique identifier for each article of Chosun Ilbo.
page_no	The page number of the article.
title	The title of the newspaper article.
regdate	The registration date of the article.
type	The type of the article.
publication_day	The day of the week when the article was published.
section	The section of the newspaper where the article is placed.
publication_date	The date when the article was published (Year-Month-Day).
completeness	The completeness of the article ("Y" indicates Yes).
body	The main text of the article.
publication_no	The publication number.
node_id	Numeric identifier associated with the article.
source_image_file	The name of the image file.
@timestamp	The timestamp indicating when the newspaper textual data was collected to the Excel file (Figure 4).
isn	The International Standard Number (ISN) associated with the article.
source_xml_file	The XML file name.
page_section	The section of the newspaper (society, general section, advertisement, politics, culture).
url	The URL to the article.
image	The URLs containing the website links to the scanned and digitized image files of the Chosun Ilbo newspaper database (used for collecting 47,777 JPG image files in this research).
sub_title	The subtitle of the article.
authors	The author(s) who wrote the newspaper article.

Table 3 Metadata definitions of Figure 4 Excel file columns.

Figure 5 47,777 image files collected from the Chosun Ilbo News Library (1920–1940) (*ChosunIlboNewsLibrary*, 2024).

recognizes a redundant object that is absent in reality. These findings highlight areas for improvement in future research. The accessibility of these FPC images has been available through the dissemination of our Excel and CSV files on the JOHD Dataverse website (Lee et al., 2024a) (Figure 6). This Excel file contains several metadata, including URLs for YOLOv5_FPC-detected images, along with their corresponding publication dates, and the file name. Researchers now have seamless access to the FPCs from the official Chosun Ilbo website by simply clicking on the URLs they want to investigate. This enables a thorough examination of the 1,040 FPC objects, and the frequencies of these objects are analyzed in Tables 4 and 5. The detected image files include previously undiscovered FPCs by previous researchers (Figure 7).



	G
nthD	URLs
0921	https://100image.chosun.com/cs-digitizing-images/1924/09/19240921/Image/news-cs-1473-19240921-e-1-04-300-b1-bw-thumb-1100.jpg
1013	https://100image.chosun.com/cs-digitizing-images/1924/10/19241013/Image/news-cs-1495-19241013-e-1-03-300-b1-bw-thumb-1100.jpg
1014	https://100image.chosun.com/cs-digitizing-images/1924/10/19241014/Image/news-cs-1496-19241014-e-1-03-300-b1-bw-thumb-1100.jpg
1015	https://100image.chosun.com/cs-digitizing-images/1924/10/19241015/Image/news-cs-1497-19241015-e-1-03-300-b1-bw-thumb-1100.jpg
1016	https://100image.chosun.com/cs-digitizing-images/1924/10/19241016/Image/news-cs-1498-19241016-e-1-03-300-b1-bw-thumb-1100.jpg
1017	https://100image.chosun.com/cs-digitizing-images/1924/10/19241017/Image/news-cs-1499-19241017-e-1-03-300-b1-bw-thumb-1100.jpg
1018	https://100image.chosun.com/cs-digitizing-images/1924/10/19241018/Image/news-cs-1500-19241018-e-1-03-300-b1-bw-thumb-1100.jpg
1019	https://100image.chosun.com/cs-digitizing-images/1924/10/19241019/Image/news-cs-1501-19241019-e-1-03-300-b1-bw-thumb-1100.jpg
1020	https://100image.chosun.com/cs-digitizing-images/1924/10/19241020/Image/news-cs-1502-19241020-e-1-03-300-b1-bw-thumb-1100.jpg
1021	https://100image.chosun.com/cs-digitizing-images/1924/10/19241021/Image/news-cs-1503-19241021-e-1-03-300-b1-bw-thumb-1100.jpg
1022	https://100image.chosun.com/cs-digitizing-images/1924/10/19241022/Image/news-cs-1504-19241022-e-1-03-300-b1-bw-thumb-1100.jpg
1023	https://100image.chosun.com/cs-digitizing-images/1924/10/19241023/Image/news-cs-1505-19241023-e-1-03-300-b1-bw-thumb-1100.jpg
1024	https://100image.chosun.com/cs-digitizing-images/1924/10/19241024/Image/news-cs-1506-19241024-e-1-03-300-b1-bw-thumb-1100.jpg
1025	https://100image.chosun.com/cs-digitizing-images/1924/10/19241025/Image/news-cs-1507-19241025-e-1-03-300-b1-bw-thumb-1100.jpg
1026	https://100image.chosun.com/cs-digitizing-images/1924/10/19241026/Image/news-cs-1508-19241026-e-1-05-300-b1-bw-thumb-1100.jpg
1027	https://100image.chosun.com/cs-digitizing-images/1924/10/19241027/Image/news-cs-1509-19241027-e-1-03-300-b1-bw-thumb-1100.jpg
1028	https://100image.chosun.com/cs-digitizing-images/1924/10/19241028/Image/news-cs-1510-19241028-e-1-03-300-b1-bw-thumb-1100.jpg
1029	https://100image.chosun.com/cs-digitizing-images/1924/10/19241029/Image/news-cs-1511-19241029-e-1-03-300-b1-bw-thumb-1100.jpg
1030	https://100image.chosun.com/cs-digitizing-images/1924/10/19241030/Image/news-cs-1512-19241030-e-1-03-300-b1-bw-thumb-1100.jpg
1031	https://100image.chosun.com/cs-digitizing-images/1924/10/19241031/Image/news-cs-1513-19241031-e-1-03-300-b1-bw-thumb-1100.jpg
1101	https://100image.chosun.com/cs-digitizing-images/1924/11/19241101/Image/news-cs-1514-19241101-e-1-09-300-b1-bw-thumb-1100.jpg
1102	https://100image.chosun.com/cs-digitizing-images/1924/11/19241102/Image/news-cs-1515-19241102-e-1-03-300-b1-bw-thumb-1100.jpg

Figure 6 Our Dataset: "Metadata for the YOLOv5_FPC Detected Images" (Lee et al., 2024a) containing the URLs (YOLOv5_FPC-detected 1035 image files; 1040 FPC objects in total), and their publication dates sourced from the Chosun Ilbo News Library (1920–1940).

INDEX	NAME OF FPC DETECTED USING THE YOLOv5_FPC MODEL (CHOSUN ILBO NEWS LIBRARY, SPANNING 1920–1940)	FREQUENCY (PER FPC)
1	Meongteongguri	726
2	Byeokchangho	126
3	Japanese Language-Written Cartoon	104
4	Baekgongsan	13
5	Dolbo and Mikki	12
6	Ttukdugiui Seollori	11
7	Chador's Adventure	9
8	Arctic Exploration	8
9	Rubber Balloon	7
10	Football Player	8
11	Makdongi and Goose	2
12	Biography of a Fool	2
13	Paengkenggun's Monkey Catching	1
14	Buffalo and Fish	1
15	Then Yes	1
16	Hobang Bridge	1
17	Ice Snack	1
18	The Evil of Alcohol	1
19	Put Your Hands Up	1
20	Better Radio	1
21	Tiger Den	1
22	Hide and Seek	1
23	Samyeong's Caramel Cartoon	1
24	Love Trees	1

Table 4 Frequency analysis of the 1040 YOLOv5_FPC-detected FPCs discovered from the Chosun Ilbo News Library (1920–1940) (*ChosunIlboNewsLibrary*, 2024).

SERIES OF "MEONGTEONGGURI" FPC	CHUNG'S RESEARCH FINDINGS OF FPCs (CHUNG, 2016)	OUR YOLOv5_FPC FINDINGS OF FPCs	DIFFERENCES (PER FPC)
Reporter Life Part 1	None	35	+35
Modern Life	None	4	+4
Social Work	50	62	+12
Heonmulkyeoji	48	55	+7
Ssutdeokdaegi	18	19	+1
Student Life	12	12	0
Ssonawatso	9	9	0
Self-sufficiency	87	86	–1
Round the World	148	147	–1
Hunger Life	50	19	–31
Dating Life	181	178	–3
Family Life	102	100	–2

Table 5 Frequency comparison of the "Meongteongguri" series.

Figure 7 Previously undiscovered FPC image data from the Chosun Ilbo News Library digital archive (*ChosunIlbo*, 2024).

Upon examining each newspaper image files, the authors were able to discern the date indicators located in the upper right corner of each published newspaper: "日" for day, "月" for month, "年" for year. These date indicators, inscribed in Japanese Kanji — which originated from Chinese characters that derived from archaic pictograms of ancient China (Jones and Aoki, 1988)—, reflect the historical dynamics and linguistic interactions among the three East Asian countries, including Korea.

3.5 DEVELOPMENT AND IMPLEMENTATION OF THE YOLOV5_FPC-DETECTOR USING THE GOOGLE COLAB PLATFORM

Furthermore, in this research, we developed the YOLOv5_FPC-Detector script on the Google Colab platform (Lee et al., 2023a), leveraging our optimized deep learning weights of the YOLOv5_FPC model; Google Colab is a globally used open-access cloud-based computational platform, available for a broader audience. This provides advantages in terms of both time and cost efficiency for the public who want to detect FPCs in their local computers. The entire script, depicted in Figures 8, 9, and 10, accompanied by detailed usage guidelines, is now accessible on the website (Lee et al., 2023a), and our GitHub repository (Lee et al., 2024b).

Figure 8 ensures the installation of necessary libraries and dependencies essential for the operation of the YOLOv5_FPC model. The weights of the YOLOv5_FPC, which have been

```
[1] # YOLOv5_FPC Detector

    import os
    import subprocess
    from google.colab import files
    import gdown
    import zipfile
    import shutil

    !git clone https://github.com/ultralytics/yolov5.git
    !pip install -U git+https://github.com/ultralytics/yolov5.git
    !pip install -r /content/yolov5/requirements.txt

    weights_url = "https://drive.google.com/file/d/13nQ-ujmDfdmssMPzDWStYydl_5IX_sYB/view?usp=share_link"
    weights_path = "/content/weight_YOLOv5_FPC.pt"
    gdown.download(weights_url, weights_path, quiet=False)

    Cloning into 'yolov5'...
    remote: Enumerating objects: 16088, done.
    remote: Counting objects: 100% (32/32), done.
    remote: Compressing objects: 100% (30/30), done.
    remote: Total 16088 (delta 10), reused 14 (delta 2), pack-reused 16056
    Receiving objects: 100% (16088/16088), 14.65 MiB | 16.43 MiB/s, done.
    Resolving deltas: 100% (11042/11042), done.
    Collecting git+https://github.com/ultralytics/yolov5.git
      Cloning https://github.com/ultralytics/yolov5.git to /tmp/pip-req-build-fouu3y23
      Running command git clone --filter=blob:none --quiet https://github.com/ultralytics/yolov5.git /tmp/pip-req-build-fouu3y23
      Resolved https://github.com/ultralytics/yolov5.git to commit 3f02fdee1d8f1a6cf18a24be3438096466367d9f
```

Figure 8 Automatic FPC detection using the weights of the YOLOv5_FPC model on Google Colab. This script imports the weights and downloads dependencies required for the automatic detection process.

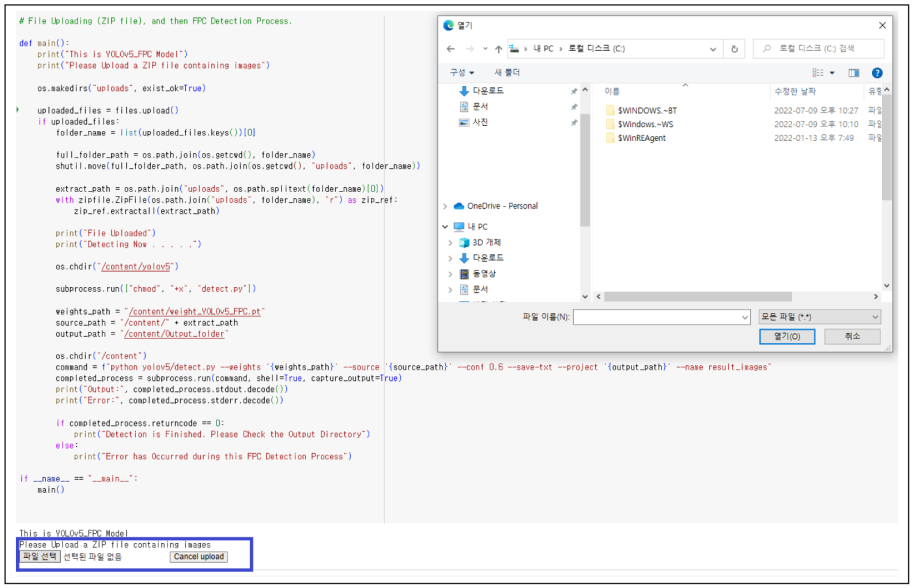


Figure 9 Users can simply upload their files to detect FPCs on their local computers.

```
# Downloading the Output File

output_path = "/content/Output_folder"

file_names_path = os.path.join(output_path, "result_images", "labels")
file_names = os.listdir(file_names_path)

source_folder = os.path.join(output_path, "result_images")
destination_folder = os.path.join(output_path, "FPC_model_detected_images")

os.makedirs(destination_folder, exist_ok=True)

for file_name in file_names:
    file_name = os.path.splitext(file_name)[0]

    if not file_name.lower().endswith(".jpg"):
        file_name = f"{file_name}.jpg"

    source_file = os.path.join(source_folder, file_name)
    destination_file = os.path.join(destination_folder, file_name)

    if os.path.exists(source_file):
        shutil.copy(source_file, destination_file)
        print(f"Copied {file_name} to {destination_folder}.")
    else:
        print(f"File {file_name} not found in {source_folder}")

shutil.make_archive("/content/FPC_model_detected_images", "zip", destination_folder)

# ZIP File Downloading Process

files.download("/content/FPC_model_detected_images.zip")
```

Figure 10 The detected FPCs are saved on their local computers.

uploaded to the Google folder, were downloaded to be utilized in this script for automatic FPC detection. Figure 9 provides a guide on how users can upload image files in ZIP format to Google Colab for FPC detection. By selecting the "Select File" option, users can simply upload their ZIP file. Once uploaded, it automatically detects FPCs within the image files they uploaded by employing the YOLOv5_FPC model. The script (in Figure 10) performs the task of aggregating these text files containing the information of the detected file names, and then cross-referencing them with the folder containing all image files. Subsequently, only those image files identified by the YOLOv5_FPC as containing FPCs are extracted, utilizing the shutil.copy() function. The resulting detected files by the YOLOv5_FPC model are saved in this directory: "/content/Output_folder/detected_true_images." To enhance user accessibility, the detected files are automatically archived into a compressed file named "FPC_model_detected.images.zip," and then saved to the user's local machine. By synergizing the YOLOv5_FPC model with the Google Colab scripts, researchers can effectively detect FPCs across various historical Korean newspapers.

4 RESULTS AND DISCUSSION

4.1 CHALLENGES AND LIMITATIONS

During our research, we encountered an issue with false positives. Occasionally, the model erroneously identified objects such as the vertically displayed Chosun Ilbo title — written in Chinese characters, and advertisements as FPCs. We believe that the model erroniously detected the vertically displayed Chinese characters and advertisements as 4 × 1 FPCs. We believe this issue arises because the vertically displayed Chosun Ilbo titles and advertisements visually resemble 4 × 1 panel FPCs. There is a need for further research to diminish the rate of these false positive detections in subsequent studies. One potential solution could be to collect more training data, specifically including 4 × 1 panel FPCs. This was challenging in our study because the vastness of Internet made locating and collecting relevant training data difficult. However, despite this obstacle, we have uncovered previously undiscovered historical FPC images (to the best of our knowledge) in the big data digital archive with this efficient computational methodology, which is now available via the JOHD Dataverse (Lee et al., 2024a), facilitating further discourse.

Our interdisciplinary research has advanced digital humanities through the application of deep learning for cultural and historical content analysis in the big data digital newspaper archive. We have elucidated how humanities researchers can enhance the achievement of humanistic goals by understanding data, algorithms, and the computational platform.

Key contributions are:

1. Collection of training data for FPCs.
2. Development of YOLOv5_FPC by fine-tuning the initial YOLOv5 model.
3. Big data collection from the Chosun Ilbo News Library (1920–1940), totaling 47,777 JPG files.
4. Data mining: Deployment of the YOLOv5_FPC model on the big data Chosun Ilbo News Library to detect FPC image objects.
5. Database curation: Uploading the Excel and CSV files containing metadata regarding YOLOv5_FPC-detected FPC image files, including previously undiscovered ones, to foster further global discourse on FPCs.
6. Data analysis of the detected FPC objects.
7. Development of the YOLOv5_FPC-Detector script, utilizing the Google Colab platform, to enhance its reuse potential.

4.2 APPLICATIONS

Future research should focus on the development of comprehensive databases containing FPCs from diverse Korean newspapers with extensive archival records using our YOLOv5_FPC model. By utilizing the YOLOv5_FPC model and the Google Colab scripts, researchers can optimally detect their FPCs from various historical Korean newspapers. Prominent among these are the Chosun Ilbo (*ChosunIlboNewsLibrary*, 2024), the Donga Ilbo (*DongaIlbo*, 2024), and the JoongAng Ilbo

(*JoongAngIlbo*, 2024) which are well-known newspaper publication companies in the Republic of Korea and contain digital archives spanning decades. This approach will enable researchers to compare and analyze historical FPCs from various publication companies. Furthermore, expanding this research methodology to include the detection of 1 × 4 matrix FPCs — which are prevalent in Western countries but not covered in this study — could enable researchers in Western countries to deepen their understanding of societal implications by investigating FPCs in future research. Researchers would be able to identify which characteristics of FPCs were prevalent during specific eras and among particular audiences, as well as how society was portrayed in the FPCs and, conversely, how the FPCs influenced society. Additionally, the adoption of Linked Research on Open Data (LOD) and Relational Database (RDB) will be essential, as they will streamline data management and improve access and connectivity of information for scholars. Such a database will be a crucial resource in cultural and historical studies, offering new insights into humanities events and under-researched areas. Moreover, the strategic application of these databases in content services could extend their utility beyond their current status quo, making them valuable tools for broader engagements. Our image detection methodology could be expanded to effectively extract diverse image content, including but not limited to, advertisements, clothing, photos, and cultural and historical image sources from large-scale databases.

FUNDING INFORMATION

This work was supported by the Korea Foundation for the Advancement of Science and Creativity (KOFAC), the Ministry of Science & ICT (Project Number: D23030005; Funding ID: N01230055), and The Chosun Ilbo Media Institute (G01230588).

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Seojoon Lee: Writing – original draft, Methodology, Data curation, Conceptualization, Software, Formal Analysis, Visualization, Writing – review & editing.

Byungjun Kim: Data curation, Validation, Conceptualization, Writing – review & editing.

Bong Gwan Jun: Project administration, Conceptualization, Writing – review & editing, Funding acquisition.

AUTHOR AFFILIATIONS

Seojoon Lee orcid.org/0009-0006-4793-6764
Graduate School of Culture Technology, Korea Advanced Institute of Science and Technology, Daejeon, the Republic of Korea

Byungjun Kim orcid.org/0000-0001-9925-9343
Center for Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology, Daejeon, the Republic of Korea

Bong Gwan Jun orcid.org/0000-0003-0900-2590
School of Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology, Daejeon, the Republic of Korea

REFERENCES

Arie, L. G. (2022). The practical guide for object detection with YOLOv5 algorithm. *Medium*. Retrieved from: https://towardsdatascience.com/the-practical-guide-for-object-detection-with-yolov5-algorithm-74c04aac4843 (last accessed: 16 April 2024).

Cao, Z., Liao, T., Song, W., Chen, Z., & **Li, C.** (2020). Detecting the shuttlecock for a badminton robot: A YOLO based approach. *Expert Systems with Applications*, *164*(8), 113833. DOI: https://doi.org/10.1016/j.eswa.2020.113833

Chen, X., Xiang, S., Liu, C.-L., & **Pan, C.-H.** (2013). Vehicle detection in satellite images by parallel deep convolutional neural networks. *2013 2nd IAPR Asian Conference on Pattern Recognition*. Naha, Japan. DOI: https://doi.org/10.1109/ACPR.2013.33

ChosunIlbo. (2024). *Chosun Ilbo*. Retrieved from: https://100image.chosun.com/cs-digitizing-images/1933/06/19330622/Image/news-cs-4325-19330622-e-1-03-300-b1-bw-thumb-1100.jpg (last accessed: 16 April 2024).

ChosunIlboNewsLibrary. (2024). *ChosunIlboNewsLibrary*. Retrieved from: https://newslibrary.chosun.com (last accessed: 16 April 2024).

Chung, H. (2016). Images of modern city, Gyeongseong seen from Manwha Fool series. *History Forum*, *43*(12). DOI: https://doi.org/10.14380/AHF.2016.43.179

Darkpgmr. (2021). Darklabel. *GitHub*. Retrieved from: https://github.com/darkpgmr/DarkLabel (last accessed: 2 March 2023).

DongaIlbo. (2024). *DongaIlbo*. Retrieved from: https://www.donga.com (last accessed: 17 April 2024).

Dwyer, B., Nelson, J., & **Solawetz, J.** (2022). Roboflow (version 1.0) [software]. Retrieved from: https://roboflow.com (last accessed: 13 January 2023).

Fiala, M. L., Nowitzki, O., Engelhardt-Nowitzki, C., & **Wöber, W.** (2022). Optical music recognition of printed white mensural notation: Conversion to modern notation using object detection mechanisms. *International Journal of Humanities and Arts Computing*, *16*(1), 33–49. DOI: https://doi.org/10.3366/ijhac.2022.0275

Girshick, R., Donahue, J., Darrell, T., & **Malik, J.** (2013). Rich feature hierarchies for accurate object detection and semantic segmentation. *arXiv*. Retrieved from: https://arxiv.org/abs/1311.2524. DOI: https://doi.org/10.1109/CVPR.2014.81

Hodel, T., Schoch, D., Schneider, C., & **Purcell, J.** (2021). General models for handwritten text recognition: Feasibility and state-of-the art. German Kurrent as an example. *Journal of Open Humanities Data*, *7*(13), 13. DOI: https://doi.org/10.5334/johd.46

Hoi, S. C. H.,Wu, X., Liu, H.,Wu, Y.,Wang, H., Xue, H., & **Wu, Q.** (2015). LOGO-Net: Largescale Deep Logo Detection and Brand Recognition with Deep Region-based Convolutional Networks. *arXiv, abs/1511.02462*. Retrieved from: https://arxiv.org/abs/1511.02462.

Jocher, G. (2020). *YOLOv5 by Ultralytics. GitHub*. Retrieved from: https://github.com/ultralytics/yolov5. (Version 7.0) (last accessed: 5 February 2023).

Jones, E. A., & **Aoki, C.** (1988). The Processing of Japanese Kana and Kanji Characters. In D. de Kerckhove & C. J. Lumsden (Eds.), *The Alphabet and the Brain*. DOI: https://doi.org/10.1007/978-3-662-01093-8

JoongAngIlbo. (2024). *JoongAngIlbo*. Retrieved from: https://www.joongang.co.kr (last accessed: 17 April 2024).

Kim, K. (2005). Reading newspapers all at once during the August 15th Liberation. *ChosunIlbo*. Retrieved from: https://www.chosun.com/site/data/htmldir/2005/11/21/2005112170320.html (last accessed: 13 November 2023).

Krizhevsky, A., Sutskever, I., & **Hinton, G. E.** (2017). Imagenet classification with Deep Convolutional Neural Networks. *Commun. ACM*, *60*(6), 84–90. DOI: https://doi.org/10.1145/3065386

Lee, S., Kim, B., & **Jun, B.** (2023a). YOLOv5_FPC-Detector Script utilizing the Google Colab platform. https://colab.research.google.com/drive/1qnCKaUGUTF5vSRdPc7DI6y7b05P8yuQ?usp=sharing (last accessed: 9 December 2023).

Lee, S., Kim, B., & **Jun, B.** (2023b). YOLOv5_FPC Train, valid, test.zip. https://drive.google.com/file/d/1F5iJ5bx-EKTlj06X5zA0yxLo1sVZBV2l/view?usp=sharing (last accessed: 19 October 2023).

Lee, S., Kim, B., & **Jun, B.** (2024a). Excel File of Metadata for YOLOv5_FPC Detected Images. *Harvard Dataverse*. DOI: https://doi.org/10.7910/DVN/DFVZWE

Lee, S., Kim, B., & **Jun, B.** (2024b). YOLOv5_FPC-Detector script. *GitHub*. Retrieved from: https://github.com/researchgitrepository/FPCResearch/blob/main/YOLOv5 FPC-Detector%20script (last accessed: 16 April 2024).

Lee, W.-S. (2005). A Study on the Rhetoric Expression in Domestic and Foreign 4 Panel Comic Strips. *Cartoon and Animation Studies, Serial No. 9*, 18–32. Retrieved from: https://koreascience.kr/article/JAKO200522941432318.page

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A., van Ginneken, B., & **Sanchez, C. I.** (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. DOI: https://doi.org/10.1016/j.media.2017.07.005

Park, K.-C. (2010). Research on the Structure of 4 Panel Comic Strips and Rhetoric Expression for the Education of Comic Contents: With of Kyunghyangsinmoon. *Cartoon and Animation Studies*, *19*, 17–35. Retrieved from: https://koreascience.kr/article/JAKO201029848353457.page

Rajeshwari, P., Abhishek, P., Srikanth, P., & **Vinod, T.** (2019). Object detection: An Overview. *International Journal of Trend in Scientific Research and Development (IJTSRD)*, *3*(3), 1663–1665. DOI: https://doi.org/10.31142/ijtsrd23422

Redmon, J., Divvala, S., Girshick, R., & **Farhadi, A.** (2016). You Only Look Once: Unified, real-time object detection. *University of Washington, Allen Institute for AI, Facebook AI Research*. DOI: https://doi.org/10.1109/CVPR.2016.91

Sa, E.-S. (2009). Development of Press Freedom in South Korea Since Japanese Colonial Rule. *Asian Culture and History*, *1*(2), 3–17. DOI: https://doi.org/10.5539/ach.v1n2p3

Siddiqui, N. (2024). Cutting the Frame: An In-Depth Look at the Hitchcock Computer Vision Dataset. *Journal of Open Humanities Data*, *10*(5), 5. DOI: https://doi.org/10.5334/johd.163

Smirnov, S., & **Eguizabal, A.** (2018). Deep learning for object detection in fine-art paintings. *2018 Metrology for Archaeology and Cultural Heritage (MetroArchaeo)*. Cassino, Italy. DOI: https://doi.org/10.1109/MetroArchaeo43810.2018.9089828

Taigman, Y., Yang, M., Ranzato, M., & **Wolf, L.** (2014). Deepface: Closing the gap to human-level performance in face verification. *2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA. DOI: https://doi.org/10.1109/CVPR.2014.220

Xiao, Y., Tian, Z., Yu, J., Zhang, W., & **Sun, K.** (2020). A review of object detection based on deep learning. *Multimedia Tools and Applications*, *79*(35), 23729–23791. DOI: https://doi.org/10.1007/s11042-020-08976-6

TO CITE THIS ARTICLE:
Lee, S., Kim, B., & Jun, B. G. (2024). Automatic Detection of Four-Panel Cartoon in Large-Scale Korean Digitized Newspapers using Deep Learning. *Journal of Open Humanities Data*, 10: 36, pp. 1–15. DOI: https://doi.org/10.5334/johd.205

Submitted: 06 March 2024
Accepted: 19 April 2024
Published: 06 June 2024

COPYRIGHT:
© 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

Journal of Open Humanities Data is a peer-reviewed open access journal published by Ubiquity Press.