

Reading Gokturkish text with the Yolo object detection algorithm

Mevlut Karakaya¹, Sadberk Ersoy², Ahmet Feyzioğlu³, Sezgin Ersoy⁴

^{1, 3, 4}Advanced Research in Mechatronics and Artificial Intelligence, Marmara University, Istanbul, Türkiye

²Technische Universität Braunschweig, Institut für Geschichtswissenschaft Geschichte und Geschichtsdidaktik, Braunschweig, Germany

⁴Technische Universität Braunschweig, Alte Salzdahlumer Str. 203, Germany

⁴Corresponding author

E-mail: ¹mevlutkarakaya1998@gmail.com, ²sadberkersoy@gmail.com,

³ahmet.feyzioglu@marmara.edu.tr, ⁴sezgin.ersoy@tu-braunschweig.de

Received 30 October 2023; accepted 4 December 2023; published online 4 January 2024

DOI <https://doi.org/10.21595/jmai.2023.23836>



Copyright © 2024 Mevlut Karakaya, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. This study has important scientific, cultural and economic contributions. From a scientific point of view, the decipherment of Gokturkish texts is of critical importance for research on Turkish culture, history and language. This study will enable historians and researchers to analyze these documents more quickly and effectively. Culturally, the reading of Gokturkish texts will help us gain a deeper understanding of Turkish culture and history. For linguists and cultural researchers, understanding these texts can offer new perspectives on the richness and cultural heritage of the past. From an economic point of view, this thesis argues that computer-assisted reading technology can contribute to a faster and more efficient reading and understanding of Gokturkish texts, making it easier to analyze the documents. This in turn frees up more time and resources for researchers and cultural experts, allowing them to focus on future work.

Keywords: YOLO, antic languages, Gokturkish, nature language processing.

1. Introduction

Natural language processing (NLP) is a branch of artificial intelligence (AI) and enables computers to comprehend [1], produce, and manipulate human language [2]. Natural language processing has the feature of querying data with real language text or voice [3]. It provides services in many areas, especially health [4], law [5], finance [6], security [7], arts [8] and education [9]. It is important to establish a language processing process [13] based on pre-processing techniques [10], programming languages [11], and library development [12]. It is very important for us to transfer the languages that are disappearing to new generations and to carry historical knowledge to the present day [14]. Through various awareness projects within the UN [16], especially UNESCO [15], efforts are being made to raise awareness about teaching native languages in schools in order to prevent the loss of languages.

Ancient Anatolian languages [17-21] and written languages constitute important historical data. Ancient languages in Anatolia had an alphabet like today's. These scripts have a variety of primitive writing systems such as Cuneiform, Aramaic, Phoenician, Phrygian, Carian, Lycian, Side, Assyrian and Greek Alphabets. With the data obtained from reading these articles, it will be possible to read and understand history better [22], [23]. In addition, the decrease in the number of people who know these writings will cause the connection between the future and the past to be severed. In this study, a study developed to read the Gokturkish language, which is at risk of being lost, on the basis of artificial intelligence and to use it in different fields, will be presented. There is no way yet to read and record it with Natural Language Processing techniques and Artificial Intelligence.

In this study, the study of experiencing and reading Anatolian Ancient Languages and Philology, History and Engineering approaches were applied in an interdisciplinary manner and the reading of Gokturkish texts, which is a part of the study, was presented.

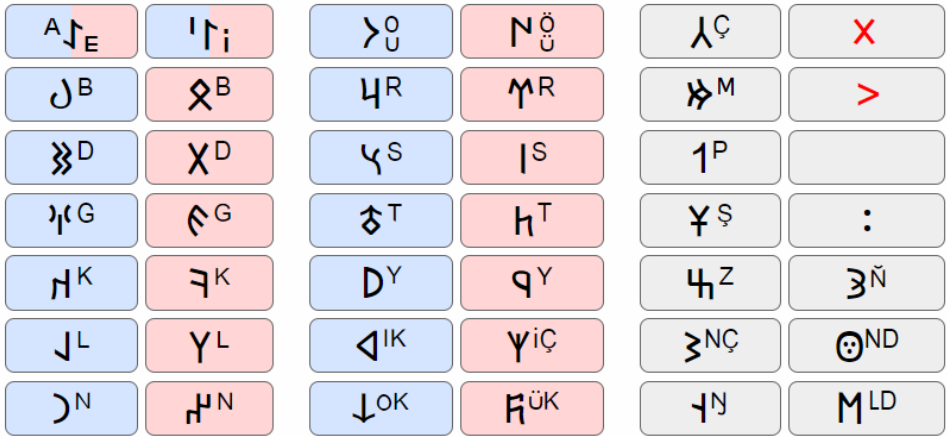


Fig. 1. Gokturkish letters [24]

2. Method

This research is deeply rooted in specific theoretical frameworks and models, with a primary emphasis on investigating the YOLO algorithm, which stands as a widely employed deep learning methodology within the expansive realm of object detection. YOLO, or You Only Look Once, is distinctive in its capability to simultaneously detect all objects present in a single image, showcasing its efficiency in real-time applications.

The investigative focus of this study extends to the utilization of pre-processed data that has been meticulously tailored to align with the unique linguistic characteristics of Gokturkish texts. Rigorous scrutiny by linguists has ensured the linguistic appropriateness of the texts for effective machine learning. Additionally, this research has benefited from consultations with subject matter experts specializing in the historical periods covered by the Gokturkish texts, contributing valuable insights to the contextual understanding of the data.

In the contemporary landscape, the applications of object detection methods have become pervasive across diverse fields, owing to continuous technological advancements and the introduction of novel architectural paradigms. The evolution of these technologies has resulted in the development of faster and more accurate object detection models, with the YOLO algorithm emerging as a standout choice for real-time object detection scenarios. Leveraging convolutional neural networks (CNN), YOLO demonstrates a high degree of efficiency and precision in the detection of objects within images. In essence, this study not only delves into the technical intricacies of the YOLO algorithm but also contextualizes its application within the broader landscape of technological progress and interdisciplinary collaboration.

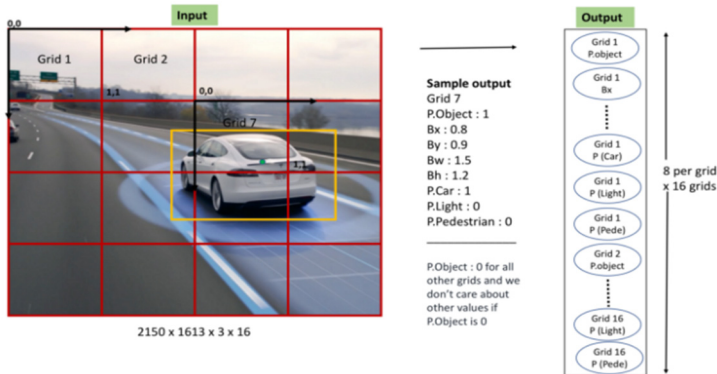


Fig. 2. YOLO working algorithm [26]

The YOLO (You Only Look Once) algorithm operates by dividing the entire image into grid regions of size $A \times A$. Each grid is then processed through a neural network to determine the presence of an object within it. If an object is detected, the algorithm identifies whether the midpoint of the object lies within the grid. Subsequently, it predicts parameters such as the object's width, length, height, class, and a confidence score.

For instance, in Fig. 2, if the midpoint of a car corresponds to the 7th grid, that particular grid is responsible for detecting the car and drawing a bounding box around it. YOLO generates a distinct prediction vector for each grid, and within each vector, the following information is included:

Confidence Score: This score indicates the model's confidence in whether an object exists in the current grid. A score of 0 signifies that the object is definitely not present, while a score of 1 indicates a high certainty of presence. This score reflects the model's confidence not only in the existence of an object but also in accurately identifying the object and determining the coordinates of the bounding box around it.

Bx: x coordinate of the midpoint of the object.

By: y coordinate of the midpoint of the object.

Bw: Width of the object.

Bh: Height of the object.

Connected Class Probability: The number of predictive values as many different classes as there are in our model.

$$\text{Confidence Score} = \text{Box Confidence Score} \times \text{Connected Class Probability}.$$
$$\text{Box Confidence Score} = P(\text{object}) \times \text{IoU}.$$

$P(\text{object})$: Probability of whether the box covers the object or not.

IoU: Intersection of the box where the object is actually located and the box that is predicted.

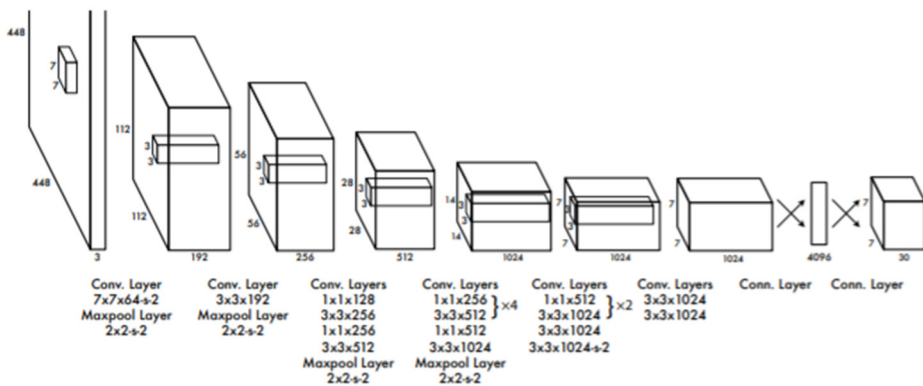


Fig. 3. Architecture of YOLO [27]

The YOLO (You Only Look Once) architecture, a groundbreaking approach in the field of object detection, draws inspiration from GoogleLeNet while introducing innovative features to address specific challenges. The YOLO architecture is characterized by a series of 24 convolutional layers designed for efficient feature extraction. Subsequently, it incorporates two fully connected layers that play a crucial role in estimating bounding box coordinates and the corresponding probabilities of detected objects.

One notable challenge in object detection, especially with Convolutional Neural Networks (CNNs), lies in the tendency to downsample input images, making it challenging to accurately recognize small objects. YOLO tackles this issue by implementing a unique strategy within its architecture. For instance, it employs a process where a layer with dimensions $28 \times 28 \times 512$ is reduced to $14 \times 14 \times 2048$, and this reduced layer is then added behind the output layer with dimensions $14 \times 14 \times 1024$.

This approach not only mitigates the limitations associated with recognizing small objects but

also enhances the network's capacity to capture intricate details and spatial relationships within the input data. The strategic reduction in layer dimensions enables YOLO to strike a balance between preserving fine-grained information and optimizing computational efficiency.

In summary, YOLO's architecture, inspired by GoogleLeNet, stands out in its incorporation of specialized layers and techniques to address challenges inherent in object detection, showcasing a commitment to both accuracy and efficiency in processing complex visual data.

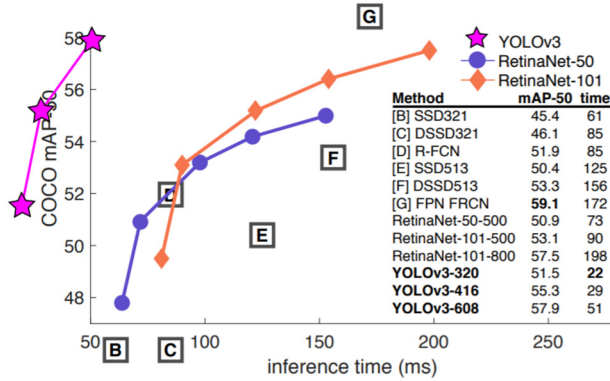


Fig. 4. YOLOv3 success chart [28]

Fig. 4 provides a comprehensive comparison between YOLOv3 and other algorithms, specifically assessing their performance at a 0.5 Intersection over Union (IoU) or mean Average Precision at 50 % overlap (mAP-50) on the COCO dataset. The graph unequivocally demonstrates YOLO's superiority over its competitors, excelling in both time efficiency and accuracy. To appreciate why other object detection algorithms lag in speed, it's instructive to delve into their underlying mechanisms.

Region-based object detection algorithms, exemplified by R-CNN, adopt a sequential methodology. They initially identify potential object regions and subsequently apply Convolutional Neural Networks (CNN) to each of these regions independently. While this approach yields commendable results, the drawback lies in the significant increase in computational operations, given that an image undergoes two distinct processes.

Despite attempts to enhance speed with subsequent iterations like Fast R-CNN and Faster R-CNN, the number of frames per second (FPS) remains notably low during both training and visual inspection. Notably, even with these advancements, the Faster R-CNN algorithm achieves only an average of 7 FPS in real-time scenarios. The unparalleled speed of the YOLO algorithm stems from its ability to predict all objects and their coordinates in a single pass through the neural network, treating object detection as a unified regression problem.

This distinctive prediction process enables YOLO to process images swiftly, establishing it as both a fast and accurate object detection algorithm. By fundamentally altering the paradigm of object detection, YOLO emerges as a pioneering solution that seamlessly combines speed and precision, making it particularly well-suited for real-time applications in the evolving landscape of computer vision.

The research hypotheses are that the results of this study will demonstrate the effectiveness of the use of the YOLO algorithm in reading Gokturkish texts and that this algorithm will serve as an example for other similar studies.

The YOLO algorithm used in this study is faster than traditional object detection methods and provides high accuracy rates. YOLO can simultaneously detect and classify objects in a single image. These features are used for example, traffic flow monitoring, facial recognition, object tracking, security systems, etc. However, the YOLO algorithm has not been used before in Turkic languages such as Turkish or Gokturkish. Therefore, in this thesis, the use of the YOLO algorithm in Gokturkish texts will be examined.

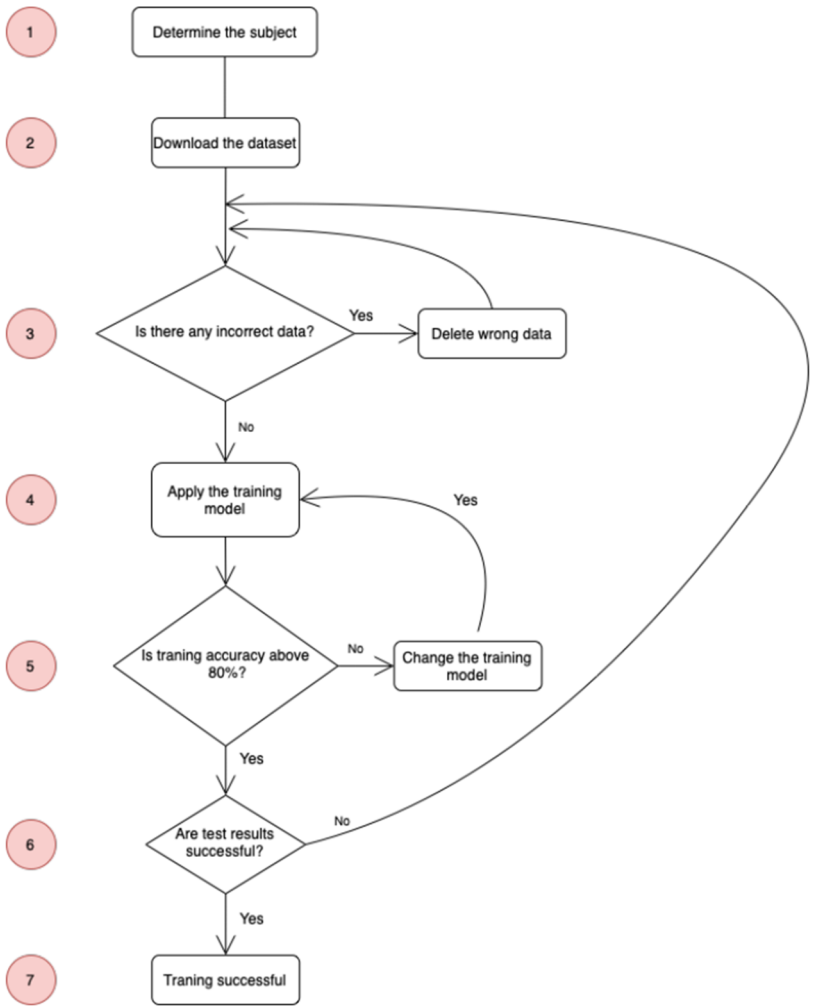


Fig. 5. The training algorithm [25]

3. Conclusions

This comprehensive study is poised to deliver multifaceted benefits across scientific, cultural, and economic domains. At its scientific core, the research endeavors to deepen our understanding of the intricate processes involved in reading and decoding Gokturkish texts, promising to catalyze further investigations in this burgeoning field of study. Culturally, the study holds the promise of enriching our comprehension of Turkish culture and history, playing a pivotal role in the preservation of invaluable information that might otherwise be lost. On an economic front, the demonstration of the YOLO algorithm's prowess in facilitating faster and more efficient reading of Gokturkish texts stands to streamline the work of historians, archaeologists, and other researchers, thereby enhancing overall productivity and effectiveness.

In essence, this study represents a concerted effort to evaluate the efficacy of employing the YOLO algorithm for the interpretation of Gokturkish texts. By fostering interdisciplinary collaboration, it aims to underscore the importance of leveraging computer-aided technology in decoding ancient languages. The potential applications of this research are far-reaching, encouraging diverse disciplines to join forces in unlocking the mysteries embedded in Gokturkish texts. Through this synergy, the study not only stands to advance our scientific understanding but

also promises to contribute significantly to the cultural and economic landscape.

The investigation into the utilization of the YOLO object detection algorithm for computer-assisted reading of Gokturkish texts holds particular significance for history, language, and cultural research. By expediting the deciphering of historical documents, this thesis has the potential to grant researchers quicker and more effective access to the wealth of information contained within these texts. Moreover, the original contribution lies in exploring the YOLO algorithm's capacity to navigate the nuances of Gokturkish texts, aiming to make a meaningful impact on the broader field of Natural Language Processing, especially in the realm of ancient or lost languages.

Rooted in expert analyses and textual readings, this study establishes a fundamental letter-word relationship, imparting this knowledge to the system. By doing so, it not only bridges the gap between traditional disciplines but also propels the study of Gokturkish texts into the forefront of cutting-edge technology applications. As a melting pot of diverse disciplinary interests, this study serves as a catalyst for collaboration, building bridges between history, language, culture, and computer science, ultimately contributing to a more holistic and nuanced understanding of Gokturkish texts.

Acknowledgements

The authors have not disclosed any funding.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

Mevlut Karakaya: contributions: data curation, funding acquisition, investigation, resources, software, validation. Sadberk Ersoy: data curation, funding acquisition, investigation, resources, software, validation. Ahmet Feyzioğlu: conceptualization, formal analysis, project administration. Sezgin Ersoy: methodology, supervision, writing – original.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] V. Raina and S. Krishnamurthy, *Building an Effective Data Science Practice*. Berkeley, CA: Apress, 2022, <https://doi.org/10.1007/978-1-4842-7419-4>
- [2] A. Dayan and A. Yilmaz, “Modelling the machines’ language with natural language processing and machine learning algorithms,” *DÜMF Mühendislik Dergisi*, Vol. 13, No. 3, pp. 467–475, Jul. 2022, <https://doi.org/10.24012/dumf.1131565>
- [3] A. Feder et al., “Causal inference in natural language processing: estimation, prediction, interpretation and beyond,” *Transactions of the Association for Computational Linguistics*, Vol. 10, pp. 1138–1158, Oct. 2022, https://doi.org/10.1162/tacl_a_00511
- [4] T. Zhang, A. M. Schoene, S. Ji, and S. Ananiadou, “Natural language processing applied to mental illness detection: a narrative review,” *NPJ Digital Medicine*, Vol. 5, No. 1, pp. 1–13, Apr. 2022, <https://doi.org/10.1038/s41746-022-00589-7>
- [5] Wang, N., Tian, and M. Y. (2023)., ““Intelligent justice”: human-centered considerations in China’s legal AI transformation,” *AI and Ethics*, Vol. 3, No. 2, pp. 349–354, 2023.

- [6] D. Biesner et al., “Anonymization of German financial documents using neural network-based language models with contextual word representations,” *International Journal of Data Science and Analytics*, Vol. 13, No. 2, pp. 151–161, Mar. 2022, <https://doi.org/10.1007/s41060-021-00285-x>
- [7] S. Salloum, T. Gaber, S. Vadera, and K. Shaalan, “A systematic literature review on phishing email detection using natural language processing techniques,” *IEEE Access*, Vol. 10, pp. 65703–65727, 2022, <https://doi.org/10.1109/access.2022.3183083>
- [8] S. Ersoy and F. Özdösemeci, “Reading and playing musical notes with image processing techniques with mobile application,” *Vibroengineering Procedia*, Vol. 44, pp. 111–116, Aug. 2022, <https://doi.org/10.21595/vp.2022.22589>
- [9] A. Ahadi, A. Singh, M. Bower, and M. Garrett, “Text mining in education-a bibliometrics-based systematic review,” *Education Sciences*, Vol. 12, No. 3, p. 210, Mar. 2022, <https://doi.org/10.3390/educsci12030210>
- [10] P. William, A. Shrivastava, P. S. Chauhan, M. Raja, S. B. Ojha, and K. Kumar, “Natural language processing implementation for sentiment analysis on tweets,” *Mobile Radio Communications and 5G Networks*, pp. 317–327, 2023, https://doi.org/10.1007/978-981-19-7982-8_26
- [11] I. Lauriola, A. Lavelli, and F. Aioli, “An introduction to deep learning in natural language processing: models, techniques, and tools,” *Neurocomputing*, Vol. 470, pp. 443–456, Jan. 2022, <https://doi.org/10.1016/j.neucom.2021.05.103>
- [12] M. S. Jahan and M. Oussalah, “A systematic review of hate speech automatic detection using natural language processing,” *Neurocomputing*, Vol. 546, p. 126232, Aug. 2023, <https://doi.org/10.1016/j.neucom.2023.126232>
- [13] A. D. Friederici, “The brain basis of language processing: from structure to function,” *Physiological Reviews*, Vol. 91, No. 4, pp. 1357–1392, Oct. 2011, <https://doi.org/10.1152/physrev.00006.2011>
- [14] G. Y. Peler, “Digital culture-2: new media-glocalization-hybridization-language-literature and folklore research,” in *Digital Culture and Language*, 2023.
- [15] “A Decade to Prevent the Disappearance of 3000 Languages,” Unesco, 2022.
- [16] “The United Nations Permanent Forum on Indigenous Issues,” 2019.
- [17] Bozgun, “Hittite tablet fragments found in Kayseri archaeological museum II,” *Archivum Anatolicum-Anatolian Archives*, Vol. 16, No. 1, pp. 51–68, 2022.
- [18] F. A. Martínez Martínez and K. V. Hernández Garay, “Nuevo posicionamiento y visibilización de la Nueva Licorera de Boyacá: Plan Estratégico de Comunicación Externa: “Primero lo Nuestro Sumercé,” Nov. 2022.
- [19] B. E. Alexandrov, “The Akkadian and Sumerian texts from Akkadian and Sumerian texts from Ortakoy-şapinuwa,” *Journal of ancient history*, Vol. 82, No. 4, pp. 983–989, 2021.
- [20] Ilker and K. O. Ç., “Old Hittite King Telipinu, his period and edict,” *Çankırı Karatekin University Karatekin Faculty of Letters Journal*, Vol. 10, No. 1, pp. 83–98, 2022.
- [21] Y. Grekyan, “Two Hurro-Urartian Lexical Parallels,” *Altorientalische Forschungen*, Vol. 49, No. 1, pp. 48–52, 2022.
- [22] A. Takahashi and H. Takahashi, “Anxiety and self-confidence in Ancient Language Studies,” 2015.
- [23] L. R. Gleitman and P. Rozin, “The structure and acquisition of reading I: Relations between orthographies and the structure of language,” in *Toward a Psychology of Reading*, 1977.
- [24] Türk Bitig, <https://www.turkbitig.com/p/gokturkishce.html>
- [25] Mevlut Karakaya, Mehmet Fatih Celebi, Akin Emrecaan Gok, and Sezgin Ersoy, “Discovery of agricultural diseases by deep learning and object detection,” *Environmental Engineering and Management Journal*, Vol. 21, No. 1, pp. 163–173, Jan. 2022.
- [26] “Evolution of object detection and localization algorithms,” Bennyilluminatedjj, 2019.
- [27] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, <https://doi.org/10.1109/cvpr.2016.91>
- [28] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” *arXiv:1804.02767*, Apr. 2018.



Mevlüt Karakaya works as a mechatronics engineer in an automation company. He has a bachelor's degree from Marmara University Mechatronics Engineering department. He continues his education as a master's student at the same university. He works on artificial intelligence and automation.



Sadberk Ersoy, graduated from the Department of Turkish Language and Literature Teaching. He completed his Thesis and Non-Thesis Master's Degrees in the same field. He is currently continuing his doctoral education. He continues his scientific studies at the Technical University of Braunschweig.



Ahmet Feyzioğlu received his B.Sc., M.Sc., and Ph.D. degrees in mechanical engineering from Marmara University. He also has a M.Sc. degree in industrial engineering from Middle East Technical University. His current research interests include mechanical design and production, system identification and control, and sustainability.



Sezgin Ersoy is a Professor of Mechatronics Engineering and Material Science. After graduating from Marmara University, he became a faculty member at the same university. His publications include a variety of efforts to understand changes in automotive mechatronics, polymer science and biomedical technologies. He is the editor and author of books. He has two science national awards and is an Editorial Member of several scientific journals.