

Advances in Object Detection Using CAX and Artificial Intelligence

Ciprian DRAGNE

[0000-0002-0666-2334], Mechanical Dept., Institute of Solid Mechanics
Romanian Academy, Street C-tin Mille 15, București, România

*Corresponding author details: PhD Ciprian DRAGNE; ciprian_dragne@yahoo.com

ABSTRACT

Object detection is a computer processing technique that uses AI to detect specific objects in digital images and videos. Techniques for the detection and recognition of objects have experienced continuous development over recent years, as their application and benefits are so very obvious. Whether they are involved in driving cars, environmental surveillance and security, assistive living for people with different disabilities, advanced robotic surgery, recognition of any types of objects (traffic signs, cars, people, fruits and vegetables, any kind of specific components, etc.). These techniques have become indispensable in research, development, and supervising the environment and such specific products. This paper proposes innovative methods for object recognition using advanced CAD tools (CAX), when the personalized database contains not images of objects, but CAD models of them, including evaluating the positioning of target objects in 3D space, obtaining position information based on advanced labeling information stored in the database, and assessing distances with these tools.

Keywords: object detection; CAX; AI.

INTRODUCTION

Object detection is even a new science today. The object detection techniques were developed in a few specific directions:

- Detection of objects and classification;
- Assessment of objects' pose and orientation;
- Assessment of object size and measurements of distances;
- Some particular object properties – color, growing stage, degradation stage, health monitoring by detection and evaluation of human skin spots, etc.;
- Tracking of movement in digitally recorded videos.

That's why object detection techniques have begun to appear massively in practice, especially used in complex assemblies like robots, surgery robots, self-driving cars, surveillance UAVs, etc. Almost all these applications use special techniques based on pre-saved dataset information. These databases must contain enough images of the object of study, from various positions, at various stages of development (in the case of fruits, vegetables, and systems that are assembled on production lines), with various specific coloring modes, etc.

The last CAD applications increase, to a great extent, the productivity of engineers or for academic study, and researchers [1]. But the development of virtual CAD models alongside experimental model control applications led to the emergence of applications that could manage both models simultaneously. Thus, the idea of digital twin applications was born and developed as a very useful tool [2].

The development of research IT applications has reached the point where many decisions are made internally by programs, without direct human intervention.

These techniques were developed with the help of a computer system to perform tasks typically similar to human intelligence, such as the way is interpreting images and vision, learning, reasoning, making decisions, problem-solving, and estimating the best next steps. That's why many of these applications are now called artificial intelligence technologies. The use of AI proves to be of real use in the faster selection, evaluation, and optimization of data types through the neural network methods specific to these artificial intelligence methods [3].

CURRENT STATE

Object detection is the very foundation of computer vision research, and involves identifying objects in images or videos and determining the position of these objects in those images using patterns to mark the areas they fall into. Object detection has become widely used in computer vision tasks, such as for image annotation [4], type vehicle detection and counting [5], activity recognition [6], facial recognition, object tracking, and tracking a person in a video [7], and even in the medical application domain [8][9].

Computer vision combined with artificial intelligence will transform the way medical diagnoses are made, surgeries are performed, prosthetics are implanted,

and patients are monitored and diagnosed [11][12].

In most situations, the stored images are still obtained randomly, from different databases, developed by various people or institutions, according to different principles and purposes, images taken under various shooting conditions, with various lenses and resolutions, etc., which makes the task of object detection significantly more difficult [13][14][15].

A separate aspect should also be mentioned here, namely, IT methods for object recognition. Since the beginning of object detection, the neural network methods have proven to be the most effective. From classical techniques to convolutional ones, with or without special segmentation methods, with or without special reorientation and scaling methods, etc., many researchers are trying to improve these techniques [16][22].

Many researchers have also developed special techniques that use multiprocessor functions based on the full power of the computer or even the outstanding performance in image processing using the GPU graphics processor [23]. Some researcher improves object detection and image interpretation with special techniques for the assessment of colors [24][25][26]. Other researchers have studied and invented methods for evaluating the orientation of objects or improving the detection of their orientation [27][35]. Some research in the field has also aimed at evaluating the sizes of detected objects and the distances between different objects [36][37][38].

All of these methods required a huge amount of study, development of IT routines, and testing in various situations [39]. Many applications involve Digital Twins development [40] and are dedicated to a specific domain of activity [41]. Some object detection applications have already been integrated into mobile phones [42] The evolution of mobile phones, predicted by me and Mark Zuckerberg, will be glasses with augmented reality, internet connectivity, and object detection included for VIPs (people with visual impairments).

METHODOLOGY

This section explains some of the methods developed during this research, which have become very complex over the years. From the initial main task of conceiving, designing, researching, developing and building a test model for a surgical robot, many other research aspects have emerged such as those in this paper, because such a complex robot has many aspects necessary to be studied and all aspects are interconnected, from general medical constraints to mechanical, practical, computer vision and software implementation constraints, and in each research chapter it was necessary to develop new specific methods.

CAX technology

What is CAX? CAX (Computer-Aided Extended) is an assembly in IT of various computer-aided technologies, a suite of IT tools that utilize concurrent engineering to solve multi-objective design problems. The old design fashion style in design is still a sequential process, but the CAX technology is a parallel management process. Figure 1 shows the GUI (Graphic User Interface) of this SOLIDWORKS Add-on for assessment of a virtual assembly. A CAX tool is at this moment State-of-Art in computer-aided technology.

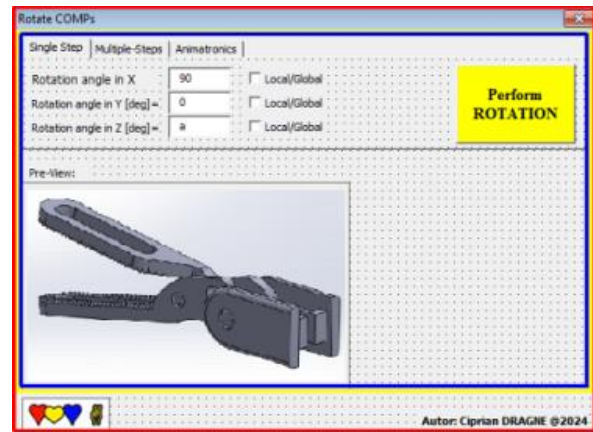


FIGURE 1: GUI of SOLIDWORKS-addon for Assembly Assessment.

The following examples show a unique SOLIDWORKS add-on that extends the capabilities of the original program for many types of assessments of a complex assembly, including control of a real robot, measurements and comparisons with a mathematical model, with AI techniques implemented inside, etc. In this particular paper, this application was used to control the CAD model in a specific view orientation, extract images shown by SOLIDWORKS, start comparisons with the target image or training with a Neural Network (NN), and detection with AI. This is a classical method for Object Detection.

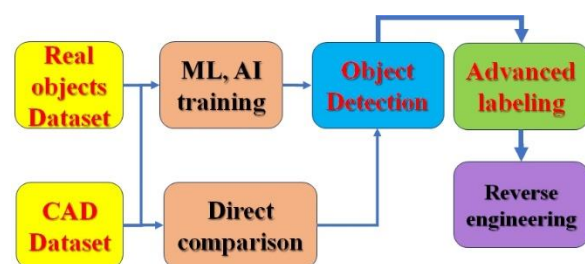


FIGURE 2: Complete Diagram of Methodology.

Figure 2 shows the complete diagram of the methodology used in this paper. Object detection techniques can use two different methods: direct comparisons and AI detection.

Object detection based on a custom CAD set

Classical method in Object Detection is to use a dataset with many images of the focused target object, train a NN, and apply an AI detection. Custom dataset is built based on CAD-Addon to make the

orientation of the virtual assembly in specific view directions. In this paper was used the direction definition based on Euler angles orientation all around 360 degrees of the CAD assembly, with a 10-degree step. This means that the custom dataset has 46656 images. There could exist some additional images if we want to save some specific assembly views. So, it is not necessary to have various images of the real objects. But the dream of any researcher working in object detection is to be able to achieve all of the previous considerations with as few images as possible of the objects being searched.

Object detection using AI and DTs

Object detection can also be evaluated by direct comparison of two images, one image of an object stored in a CAD model and the other target image. All data is stored temporarily only in computer memory. The development of IT applications for object detection research in CAD or more advanced CAX-type applications in applied engineering, and the observation of objects according to multiple selection criteria, multiple viewing directions, has brought enormous benefits to the development of this particular engineering science.

Advanced labeling

Classical Object Detection methods show only the class of the object and the confidence value in detection. My methods also show additional information stored in the dataset together with the results and the orientation information obtained after detection, which reveal the most probable 2D or 3D orientation of the object observed in the target image.

Reverse engineering in Object Detection

The Classical Object Detection method has a dataset with images stored somewhere on a computer or the internet. My methods have stored CAD models of the focused objects. But a CAD model is a parametric virtual model. So, this tool can also be used to make an update and correlation of the CAD model with an image of a real object. This is a Reverse Engineering methodology that can extend the capabilities of the initial method.

CASE STUDIES

In this section, examples of how this complex DTs application was used in Object Detection will be presented.

Direct comparisons in DTs application

For direct comparison, the SSIM criterion was used to search for the best fit between images. Images obtained from CAX application and CAD dataset can be stored temporarily in a local folder by the DT application and deleted after detection, with only the detection results being retained. So, a dataset with many images is not required if we want. But, still, CAX application can create a complete dataset of CAD viewing positions images if we want to have finally also images for a pre-trained archive to be exported for other AI detection applications and for faster detection.

Figure 3 shows the results of detection based on direct comparisons, where obtained a SSIM=0.72 (Figure3.a) for a simple object. Figure 3.b shows the best fit orientation of CAD assembly detected by CAX application.

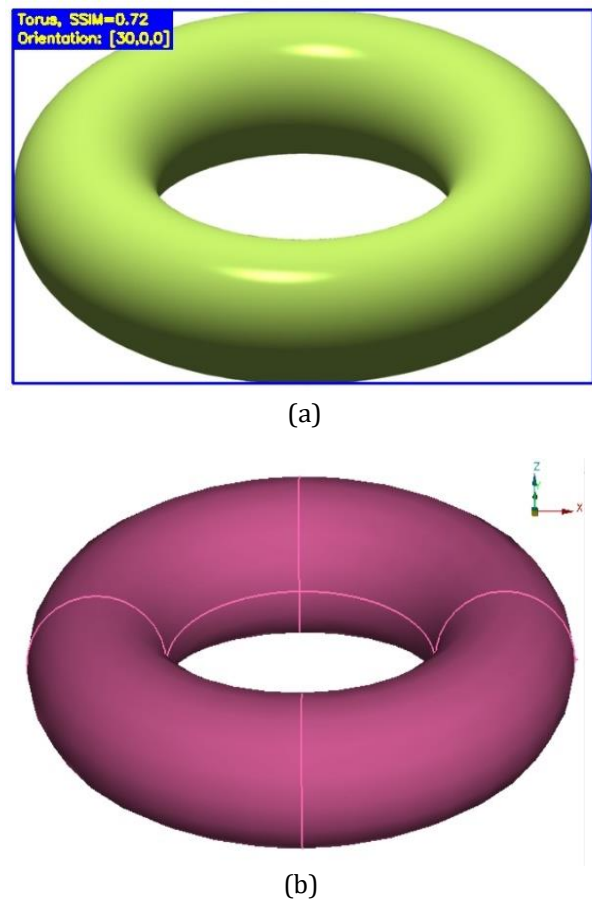


FIGURE 3: Detection Based Direct Comparison.

It is not necessary to initially store any images on your computer for any images-set, only the target object. Another example for test is presented in Figure 4 to search for a different view angle.

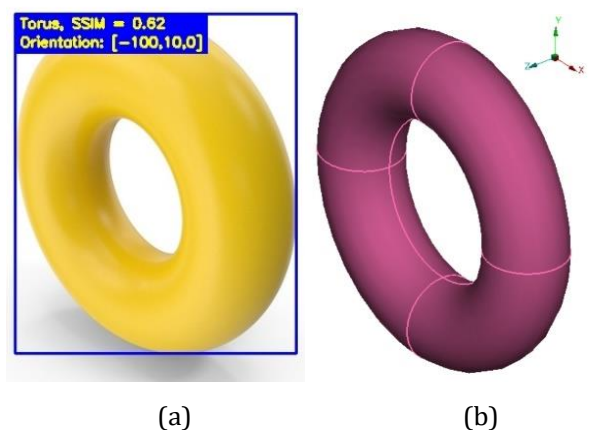
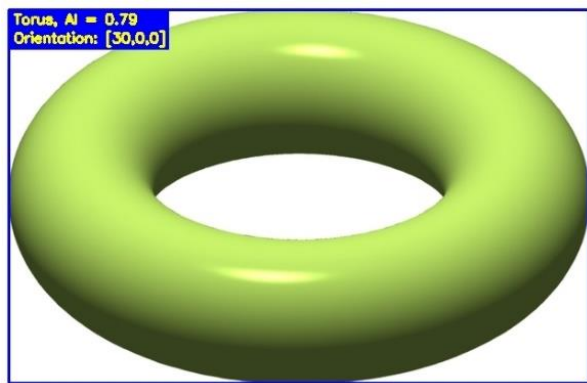


FIGURE 4: Detection Based Direct Comparison.

Object detection using DTs and AI

AI detection uses a dataset saved by the CAX application with images of the CAD model at various viewing angles. Object and orientation detection-based AI training are shown in Figure 5.



(a)



(b)

FIGURE 5: Detection Based AI Training.

Complex assemblies

In the next subchapter was tested also a complex assembly for the same method using AI. The images from dataset were obtained with CAX application.



(a)



(b)

FIGURE 6: Object Detection of Complex Assembly. (a) Images Set From CAD-Set. (b) Detection Results Using AI.

With the help of the CAX application, you can also obtain the orientation in plan (2D) of objects based on CAD models. The Figure 7 shows a such an assessment.

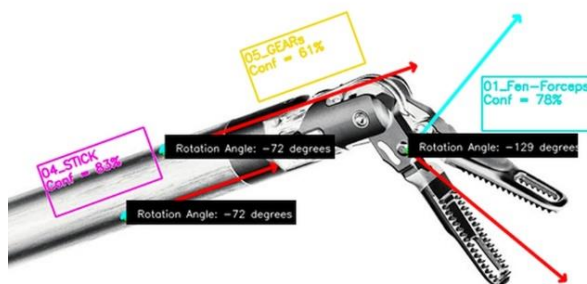


FIGURE 7: Detection of In-Plane Orientation.

This technique was developed based on YOLO11 detection methods. Since in many AI pre-trained databases, images are increased in number by rotating, scaling, and cropping 2D views, we also put some results in Figure 8 as evidence of detection for such photos as well.

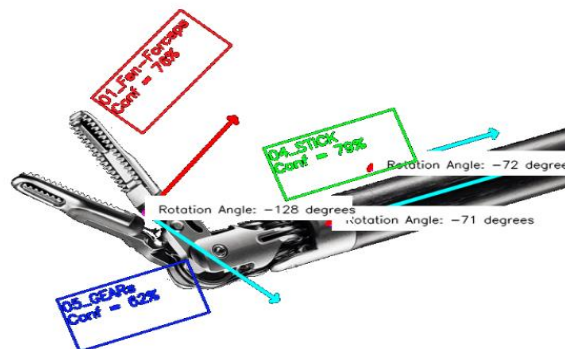


FIGURE 8: Detection of In-Plane Orientation of Rotated Images.

Figure 9 shows the detection with orientation in 2D and 3D space for a CAD detection.

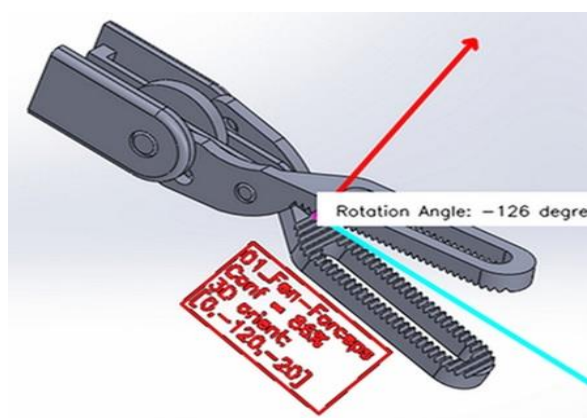


FIGURE 9: Detection of Orientation in 2d and 3d Space.

The detection in the same image from Figure 9, but rotated in the plane, is shown in Figure 10.

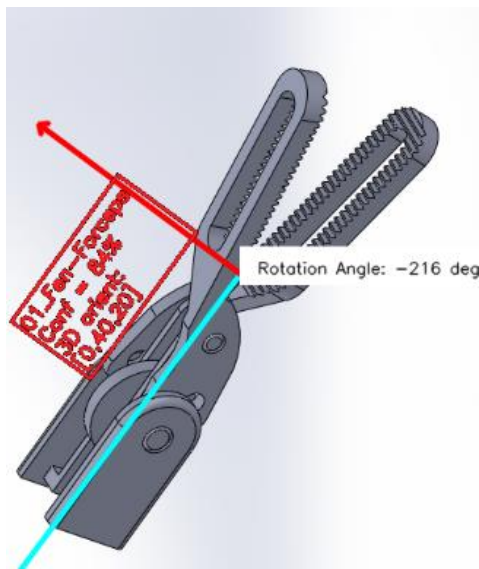


FIGURE 10: Detection of Orientation in 2D and 3D Space for Rotated Image.

Various detections for other poses of the complex assembly can be viewed in Figure 11. Home position of the assembly is presented in Figure 11 (a and b) with direct comparison and AI detection. Figure 9-c shows a rotated view assembly by the Y-axis with 40 degrees, and Figure 11-d a rotated view assembly by 20 degrees in Y Y-axis and 30 degrees in Z Z-axis.

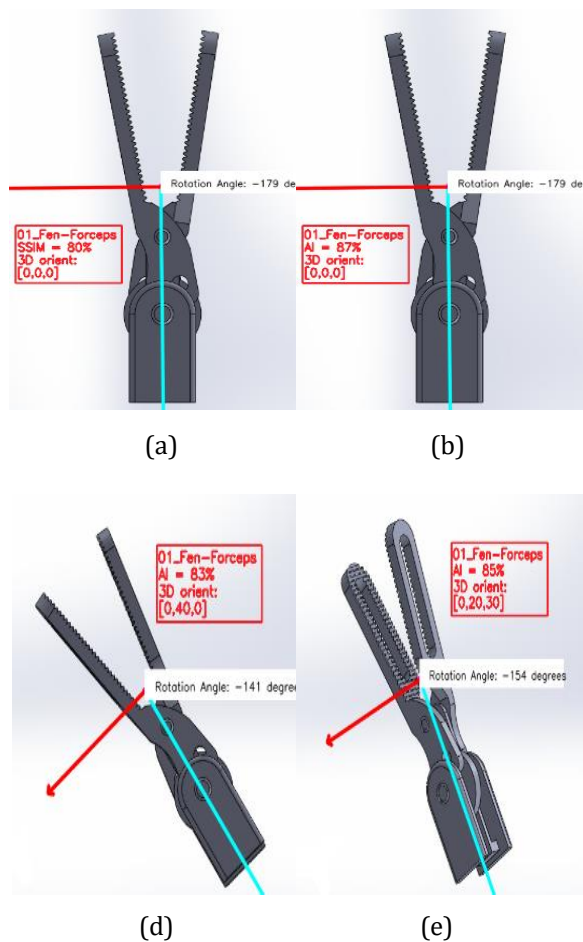


FIGURE 11: Various Assembly Pose and Detection Results.

Advantages of using CAX and AI:

- CAX uses the full performance of the professional CAD software (parametric design, smart-assemblies, parts colorization, cutting sections, transparency, animatronics, viewing angles, specific colorization, pattern shading, etc.);
- Once AI model was built, detection takes small time. Instead, for the sequential classic method is necessary to check entire data-set every time is used;
- ANN provide not only confidence values related to comparison in dataset, but also multiple type of data implemented in advanced labelling (method used, orientation, etc.);
- Images dataset for AI detection was automatically generated by CAX application.

Disadvantages of using AI and Concurrent Engineering:

- Method implemented in CAX are difficult for most researchers;
- CAX methods needed to be developed by users;
- Are necessary professional skills from multiple domains in activity: mathematics, engineering, programming, data measurements, CAD, other programming languages, etc.;
- ANN model is not small, in our case was used 46656 images, but automatically generated by the CAX application;

A video presentation with many research results, including for computer vision, similar to this paper, is also presented in this web link [47].

The current research also included the development of applications for a wide range of concurrent engineering functions. The method includes:

- CAD - CAX type tools;
- Automatic bounding-box detection and cropping image techniques are implemented;
- Tools for database management and fully automatic database construction,
- AI techniques based on programs such as PYTHON or C++ codes, ...
- Tools for storing models and machine learning (ML),
- Tools for training neural models,
- Tools for object detection and result extraction, and...
- Tools for visualizing results, along with much richer information labels than in classic AI programs.

Validation of the CAD-dataset can also be done using a real image dataset. By reverse engineering from real images to CAD-set, and from CAD-set to images again, with CAD-addon we can improve research evaluation of object detection, classification, segmentation, or orientation.

Disadvantages of custom detection methods:

- The method for object orientation has some glitches in detection for images with more than one object. Some issues can be solved by individual detection and research;

CAX and AI can also be used in the assessment of measurement for distances, with many successes then current methods because CAD assemblies have very known dimensions. These applications will be presented in a future paper.

Interest in Future Research:

- Implement 3D detection for videos (images in movement);
- Auto-adaptive tool for Object Detection;
- Test similarity also for other criteria;
- Applications for Manufacturing Line;
- Detection in images with more objects;
- Still needed research for the correlation between 2D and 3D orientation, because each method is based on different techniques.

CONCLUSIONS

The keys to success are the concurrent engineering applications.

Few global conclusions:

- All method presented here shows good result. Most of them are innovative methods;
- Because of symmetry of some assemblies, some viewing directions are obviously redundant. For complex assemblies, this issue doesn't exist;
- Computer-aided mechanical design with extended capabilities (CAX) will evolve in general, to new levels for integration of many tools;
- Using CAD and IT applications to their maximum capabilities can help enormously in solving many problems in mechanics, kinematics, or more efficiently any kind of scientific research;
- Artificial intelligence detection offers increased confidence over the direct detection method;
- AI is a research tool and an opportunity for humanity if we are going to use it in a proper manner.

Such a combination between a DTs, a complex CAX tool with many extensions, and the real model is a complex AI-agent and will be an important step in development under Industry-5.0 requirements.

For me, Industry 5.0 is already here, in everything I study:

- Research + Concept + Engineering + Evaluations + Optimizations + Recycling + Sustainability;
- Advanced materials and meta-materials;
- Advanced competitive engineering applications, CAX technology, AI, intelligent optimization methods;
- Updating and correlation of mathematical models with the experimental model;
- Connectivity with all subjects, advanced CAD models, and intelligent production systems;
- Animatronics, etc.

The technological leap Industry 5.0 will bring many changes in the way we do research, production, including Education. It will completely change the way we learn, approach education, understand complex phenomena, present results, discover new valences and methods, etc.

The complete task of this research was developed based on SOLIDWORKS Educational, PYTHON software, MATLAB software, YOLO tool, other internet resources [43][45], and user-defined programming routines [46].

This paper presents only the important aspects of our research, with very valuable information related to how can use a DT with superpowers for Object Detection. It is obvious that creativity will partially shift from human abilities to those of computing machines.

ACKNOWLEDGMENT



The Institute of Solid Mechanics - Romanian Academy, funded some of the doctoral and post-doctoral research from this paper.

REFERENCES

- [1] Kyratsis, P., Kakoulis, K., & Markopoulos, A.P., Advances in CAD/CAM/CAE Technologies. Machines, 8(1),13, 2020.
- [2] Dragne, C., 2025, SUPER DIGITAL-TWINS FOR ACADEMIC RESEARCH, International Journal of Scientific Advances, ISSN: 2708-7972., <https://www.ijscia.com/super-digital-twins-for-academic-research/>
- [3] Guan, L. (Ed.). (2017). Multimedia image and video processing. CRC press.
- [4] Oza, P., Sindagi, V. A., & Patel, V. M. (2023). Unsupervised domain adaptation of object detectors: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(6), 4018-4040.
- [5] A. Alsanabani, A., A. Ahmed, M., & M. Al Smadi, A. (2020, December). Vehicle counting using detecting-tracking combinations: A comparative analysis. In Proceedings of the 2020 4th International Conference on Video and Image Processing (pp. 48-54).
- [6] Wu, J., Osuntogun, A., Choudhury, T., Philipose, M., & Rehg, J. M. (2007, October). A scalable approach to activity recognition based on object use. In 2007 IEEE 11th international conference on computer vision (pp. 1-8). IEEE.
- [7] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (Vol. 1, pp. 886-893). IEEE.
- [8] Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nat. Med. 25, 44-56.

- [9] Litjens, G. et al. (2017). A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88.
- [10] Hosny, A., Parmar, C., Quackenbush, J. et al. (2018), Artificial intelligence in radiology. *Nat Rev Cancer* 18, 500–510, <https://doi.org/10.1038/s41568-018-0016-5>.
- [11] Zaidi, D. (2017), AI is transforming medical diagnosis, prosthetics, and vision aids. In *Venture Beat* <https://venturebeat.com/ai/ai-is-transforming-medical-diagnosis-prosthetics-and-vision-aids/>
- [12] Singh, R. et al. Deep learning in chest radiography: detection of findings and presence of change. *PLoS ONE* 13, e0204155 (2018).
- [13] Ross, C. & Swetlitz, I. (2018). IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments; internal documents show. In *Stat News* <https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/>.
- [14] Bamba, Y., Ogawa, S., Itabashi, M., Kameoka, S., Okamoto, T., & Yamamoto, M. (2021). Automated recognition of objects and types of forceps in surgical images using deep learning. *Scientific Reports*, 11(1), 22571.
- [15] Tan, M., Pang, R., & Le, Q. V. (2020). Efficient det: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10781-10790).
- [16] Chollet, F. (2017). Xception: Deep learning with depth wise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1251-1258).
- [17] He, K., Girshick, R., & Dollár, P. (2019). Rethinking imageNET pre-training. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 4918-4927).
- [18] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
- [19] Jegham, N., Koh, C. Y., Abdelatti, M., & Hendawi, A. (2024). Evaluating the evolution of yolo (you only look once) models: A comprehensive benchmark study of yolo11 and its predecessors. *arXiv preprint arXiv:2411.00201*.
- [20] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
- [21] Tan, M., & Le, Q. (2019, May). EfficientNET: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105-6114). PMLR.
- [22] Zhou, P., Ni, B., Geng, C., Hu, J., & Xu, Y. (2018). Scale-transferrable object detection. In *proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 528-537).
- [23] Huang, R., Pedoeem, J., & Chen, C. (2018, December). YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers. In *2018 IEEE international conference on big data (big data)* (pp. 2503-2510). IEEE.
- [24] Agossou, B.E., Pedersen, M., Raja, K, Vats, A. and Floor, P.A., (2025). Influence of color correction on pathology detection in Capsule Endoscopy. *arXiv preprint arXiv:2502.00076*.
- [25] Lin, T.Y., Goyal, P., Girshick, R., He, K. and Dollár, P., (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988).
- [26] Watine, L., Floor, P.A., Pedersen, M., Nussbaum, P., Ahmad, B. and Hovde, Ø., (2023, September). Enhancement of colour reproduction for capsule endoscopy images. In *2023 11th European Workshop on Visual Information Processing (EUVIP)* (pp. 1-6). IEEE.
- [27] Tian, J., Jin, Q., Wang, Y., Yang, J., Zhang, S. and Sun, D., (2024). Performance analysis of deep learning-based object detection algorithms on COCO benchmark: a comparative study. *Journal of Engineering and Applied Science*, 71(1), p.76.
- [28] Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., ... & Lin, D. (2019). MM-Detection: Open MMLab detection toolbox and benchmark. *arXiv preprint arXiv:1906.07155*.
- [29] Yao, Y., Cheng, G., Wang, G., Li, S., Zhou, P., Xie, X., & Han, J. (2022). On improving bounding box representations for oriented object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1-11.
- [30] Wang, K., Wang, Z., Li, Z., Su, A., Teng, X., Liu, M., & Yu, Q. (2023). Oriented object detection in optical remote sensing images using deep learning: A survey. *arXiv pre-print arXiv:2302.10473*.
- [31] Yi, J., Wu, P., Liu, B., Huang, Q., Qu, H., & Metaxas, D. (2021). Oriented object detection in aerial images with box boundary-aware vectors. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 2150-2159)

- [32] Reyes-Hernández, J. C., Alomar, A., Rubio, R., Piella, G., & Sukno, F. (2024, July). OBBaby-Face: Oriented Bounding Box for Infant Face Detection. In International Conference on Deep Learning Theory and Applications (pp. 336-350). Cham: Springer Nature Switzerland.
- [33] Ding, S., Mannan, M. A., & Poo, A. N. (2004). Oriented bounding box and octree based global interference detection in 5-axis machining of free-form surfaces. *Computer-Aided Design*, 36(13), 1281-1294.
- [34] Chang, J. W., Wang, W., & Kim, M. S. (2010). Efficient collision detection using a dual OBB-sphere bounding volume hierarchy. *Computer-Aided Design*, 42(1), 50-57.
- [35] Hoang, D. C., Chen, L. C., & Nguyen, T. H. (2016). Sub-OBB based object recognition and localization algorithm using range images. *Measurement Science and Technology*, 28(2), 025401.
- [36] Dragne C. et al. (2022). Distance Assessment by Object Detection for Visually Impaired Assistive Mechatronic System, *Mechatronics and Robotics in Vehicle and Medical Applications Jr.*, ISSN 2076-3417, MDPI Appl. Sci. 2022, <https://www.mdpi.com/2076-3417/12/13/6342>., DOI: 10.3390/app12136342 WOS.
- [37] Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003, November). Multiscale structural similarity for image quality assessment. In *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003 (Vol. 2, pp. 1398-1402). IEEE.
- [38] Sampat, M. P., Wang, Z., Gupta, S., Bovik, A. C., & Markey, M. K. (2009). Complex wavelet structural similarity: A new image similarity index. *IEEE transactions on image processing*, 18(11), 2385-2401.
- [39] Dragne, C. (2023). CAX Software - The Next Level in Computer Aided Technology, *International Journal of Advances in Engineering and Management - IJAEM*, Volume 5, Issue 7, July 2023, Page No: 551-561, DOI: 10.35629/5252-0507551561.
- [40] Dragne, C., 2024. Advanced Concurrent Engineering using CAD software - Classic, Avant-garde and AI. *International Journal of Advanced Multidisciplinary Research and Studies*, July 17, 2024, DOI: <https://doi.org/10.62225/2583049X.2024.4.4.3038>
- [41] Dragne C., Buliga T., (2025), AI in Recognition of Surgery Tools, in *Scientific Research Publishing*, <https://www.scirp.org/journal/jcc>.
- [42] <https://levelup.gitconnected.com/game-of-yolo-realtime-object-detection-on-your-phone-f0b6ac4bc75c> (accessed on September 10, 2025).
- [43] <https://www.ibm.com/topics/artificial-intelligence> (accessed on August 15, 2025).
- [44] <https://www.britannica.com/technology/artificial-intelligence> (accessed on August 10, 2025).
- [45] <https://builtin.com/artificial-intelligence> (accessed on July 12, 2025).
- [46] GRABCAD author codes, <https://grabcad.com/dr.eng.ciprian.dragne-1>, (last accessed August 10, 2025)
- [47] Video presentation: https://odysee.com/@dream_in_colors:2/ALL_Animatii_TEZA_Sistem_robotic_de_brahiterapie_cuSUBT:9