Computer Vision in Effective Education

1st Pankaj Balotra
Department of Data Science
University of Europe
Applied Sciences
Potsdam, Germany
pankaj.balotra@ue-germany.de

2nd Dr.Raja Hashim Ali
Department of Data Science)
University of Europe
Applied Sciences
Potsdam, Germany
hashim.ali@ue-germany.de

3rd Dr.Iftikhar Ahmed
Department of Data Science)
University of Europe
Applied Sciences
Potsdam, Germany
iftikhar.ahmed@ue-germany.de

***Abstract*—Computer vision holds immense potential in transforming educational practices by leveraging innovative technologies. This paper explores the integration of computer vision techniques into educational settings to enhance learning outcomes and engagement by extending the application of existing computer vision algorithms. Our methodology extends the application of existing computer vision algorithms, initially used for object size measurement from images, to the creation of interactive 3D animations in educational content. The previous approach limits till 2d image. Furthermore, from this point our research continues with an effective approach to merge blender (a 3d modeling tool) application with existing computer vision algorithm based on pixels calculation. Our methodology integrates object detection algorithms used for measuring object size from images with 3D animation tools to generate interactive educational content. By leveraging pixel calculations and measurements obtained from the object detection process, we translate this data into immersive 3D animations, extending the application of computer vision algorithms in educational settings. Our results showcase the effectiveness of integrating object detection techniques with 3D animation tools for educational content creation. While Test 1 demonstrates the success of our approach, Test 2 highlights areas for improvement, particularly in handling more complex objects. Overall, our methodology offers implications for enhancing educational practices, paving the way for future advancements in educational technology and offers advance learning environment.**

Our study presents a novel methodology that seamlessly integrates computer vision techniques with 3D animation tools, offering a transformative approach to educational content creation and paving the way for enhanced learning experiences.

***Index Terms*—Computer Vision, Education, Object Detection, 3D Animation, Interactive Learning**

I. Introduction

Education stands as a cornerstone in sculpting the trajectory of society, wielding immense influence over the future landscape. As technological advancements continue to burgeon, traditional educational methodologies undergo a metamorphosis, propelled by the potential of innovative tools and approaches. Among these, computer vision emerges as a potent force, nestled within the realm of artificial intelligence, offering a myriad of applications across diverse domains, with education being no exception. The fusion of computer vision techniques with pedagogical practices heralds a new era, where interactive and immersive learning experiences become not only feasible but also tailored to accommodate a spectrum of learning styles and preferences.

In the educational landscape, the integration of computer vision techniques transcends mere augmentation, offering a paradigm shift in how knowledge is imparted and absorbed. With its ability to discern and analyze visual content, computer vision enriches learning environments by infusing them with interactive elements and immersive stimuli. By harnessing the power of computer vision, educators can transcend the limitations of traditional teaching methods, fostering engagement and comprehension through personalized, adaptive learning experiences. Whether through real-time object recognition, gesture-based interactions, or augmented reality overlays, computer vision breathes life into educational content, transcending the confines of conventional pedagogy.

Moreover, the symbiotic relationship between computer vision and education catalyzes inclusivity and accessibility, democratizing learning opportunities on a global scale. By leveraging technology to accommodate diverse learning styles and preferences, educators can bridge gaps in comprehension and empower learners from disparate backgrounds. From providing tailored feedback and personalized learning pathways to facilitating remote education initiatives, computer vision serves as a catalyst for educational equity and inclusivity. As society navigates the digital frontier, the fusion of computer vision and education stands as a testament to the transformative power of technology in shaping the future of learning and societal progress. Previous research [1] [2] [3] [4] at the intersection of computer vision and education has been focused on enhancing learning outcomes through the utilization of advanced techniques and methodologies. One prominent area of investigation involves the widespread application of object detection algorithms, which serve as powerful tools for identifying and analyzing visual content within educational materials. By leveraging these algorithms, researchers and educators gain invaluable insights into the educational process, enabling them to uncover patterns, correlations, and underlying dynamics that contribute to effective learning experiences. Through the systematic analysis of visual content, object detection algorithms facilitate the development of tailored instructional strategies, content personalization, and adaptive learning systems, ultimately enhancing the overall efficacy and impact of educational interventions.

II. Literature Review

This literature review provide a clear vision of four major technique used in the research:

1) Object Detection and Parameter Extraction: Object detection algorithms are utilized to identify and locate objects within digital images, followed by parameter extraction to quantify their dimensions. This technique leverages computer vision algorithms such as contour detection, sorting, and pixel distance measurement. The application of OpenCV, SciPy, and numpy libraries in Python facilitates the implementation of these techniques as refer [5], [6], [7], and [8].
2) Blender Animation Generation: Blender's robust features for 3D modeling and animation are harnessed to create immersive educational materials. This involves integrating object detection results with Blender's bpy module to dynamically configure objects within the animation environment. The animation generation process encompasses scene setup, object incorporation, textual annotations, camera [9], [9], [10] and lighting arrangement, and rendering, all orchestrated through Python scripting .
3) Image Pre-processing: Pre-processing techniques are applied to enhance the quality of input images for subsequent analysis. This includes conversion to grayscale, Gaussian blur to reduce noise, Canny edge detection for edge identification, and morphological operations for boundary refinement. These steps ensure improved accuracy and reliability in object detection and segmentation as refer [1], [2], [11], and [12].
4) Integration Process Overview: The integration process bridges the gap between object detection results obtained through Python and Blender's animation environment. This involves data input from CSV files containing object parameters, scene setup within Blender, object incorporation with dynamic dimension configuration, textual annotations, camera and lighting arrangement, and rendering. The seamless adaptation of Python scripting within Blender bpy facilitates customization for diverse educational scenarios and ensures the creation of visually engaging and didactically enriching animations.

III. Our Contribution

A. Gap Analysis

Despite the significant advancements in both computer vision and education, there persists a noticeable gap in seamlessly integrating these technologies into educational practices. Existing approaches often fall short in combining real-time object detection with interactive 3D animation, thereby limiting their efficacy in delivering truly engaging educational content. Our study endeavors to bridge this gap by presenting a novel approach that harmoniously merges real-time object detection algorithms with interactive 3D animation tools. Through this integration, we aim to offer students a transformative learning experience that goes beyond traditional teaching methods, fostering higher levels of engagement and deeper comprehension of educational materials. This innovative approach represents a substantial leap forward in educational technology, with the potential to revolutionize teaching and learning practices in the digital age. By leveraging the synergy between computer vision and education, we envision a future where learning becomes more immersive, interactive, and accessible to learners of all backgrounds and abilities.

B. Research Questions

The main research questions and contributions are as follows: RQ1: How can computer vision techniques be leveraged to enhance educational content creation?

This research question explores the potential of computer vision technologies in revolutionizing educational content creation. By integrating advanced object detection and parameter extraction methods with Blender, a 3D modeling and animation platform, the study aims to develop dynamic and engaging educational materials. The contribution of addressing this question lies in demonstrating a novel approach that combines real-time object detection algorithms with interactive 3D animation tools, offering a transformative learning experience for students and enhancing their engagement and comprehension of educational materials.

RQ2: How can we extend the application of existing computer vision algorithms?

This question delves into the adaptability and extendibility of current computer vision algorithms for educational purposes. The study extends the use of these algorithms by incorporating them into an educational context, specifically for creating 3D animations based on detected object dimensions. The novelty here is in applying these technologies in a new domain, demonstrating their flexibility and effectiveness in educational content creation. The contribution and novelty of this work lie in its pioneering integration of computer vision techniques with 3D animation to create educational content. This approach not only enhances the learning experience by making it more interactive and engaging but also demonstrates the practical applicability of computer vision in education, paving the way for future research and development in this field. By addressing these research questions, the thesis contributes to the ongoing efforts to innovate teaching methodologies and enhance the learning environment through technological integration.

RQ3: What are the potential implications of integrating computer vision and 3D animation technologies in educational settings?

This question seeks to explore the broader implications of integrating computer vision and 3D animation technologies within educational contexts. By examining the potential benefits and challenges associated with this integration, the study aims to provide insights into how such technologies can enhance teaching and learning experiences. Additionally, it investigates the implications for educators, students, and

TABLE I
LITERATURE REVIEW TABLE SHOWING THE CONTRIBUTIONS OF VARIOUS AUTHORS FOR COMPUTER VISION APPLICATIONS.

Technique/Method	Existing Literature Characteristics	Novelty of Proposed Work	Dataset Used	Method(s) Used
Object Detection and Parameter Extraction	Existing literature emphasizes the application of object detection algorithms for various purposes, including surveillance, autonomous vehicles, and image recognition. The novelty lies in integrating these algorithms with educational content creation, specifically for 3D animation modeling, to enhance learning experiences through interactive visualizations [5], [6], [7], and [8].	Utilization of object detection results as parameters for generating educational animations, bridging the gap between computer vision and 3D animation for learning purposes.	Real time Image	Edge detection for image segmentation
Blender Animation Generation	Previous research extensively explores Blender's capabilities for 3D modeling and animation in fields such as film production and game development. The innovation here lies in integrating Blender with object detection outcomes to create dynamic educational animations, offering educators and content creators a powerful tool for visualizing complex concepts [9], [9], [10].	Integration of object detection results with Blender bpy for educational animation generation, facilitating tailored animations to elucidate intricate concepts and enhance comprehension.	Coustumized Image Data set	Image analysis and computer vision techniques
Image Pre-processing	The existing literature discusses various image pre-processing techniques to improve the quality of input images for subsequent analysis. The novelty in this work is the application of these techniques specifically for enhancing object detection and segmentation accuracy, thereby laying the groundwork for precise parameter extraction and animation generation ([1], [2], [11], and [12]).	Implementation of image pre-processing techniques to enhance object detection accuracy, ensuring reliable parameter extraction for immersive educational animations.	Various datasets (not specified)	YOLO-based object detection

educational institutions in terms of curriculum development, pedagogical approaches, and learning outcomes.

C. Problem Statement

The primary problem addressed in our study is the imperative need for innovative educational methodologies that effectively utilize state-of-the-art technologies to elevate learning experiences . Traditional educational approaches often struggle to engage and cater to the diverse learning preferences of modern students, leading to a disconnect between conventional teaching methods and contemporary educational needs. Our research aims to bridge this gap by proposing a novel approach that seamlessly integrates cutting-edge tools like computer vision and 3D animation into educational practices. By leveraging these technologies, we seek to revolutionize the educational landscape, providing students with immersive and interactive learning experiences that foster deeper understanding and engagement.

D. Novelty of this study

Our study introduces a groundbreaking methodology that marries real-time object detection algorithms with interactive 3D animation tools to produce captivating and educational content. The uniqueness of our approach lies in its seamless integration of cutting-edge computer vision techniques into educational settings, promising a paradigm shift in the way students engage with learning materials. Addressing the identified gaps in traditional educational practices, our study brings forth several key innovations:

- Interactive Learning Experiences: By leveraging real-time object detection and 3D animation, our approach offers students immersive and interactive learning experiences, enabling them to actively engage with educational content in a dynamic virtual environment.
- Enhanced Visual Understanding: Through the integration of computer vision techniques, students gain a deeper understanding of complex concepts by visually interacting with 3D representations of educational material, fostering spatial awareness and conceptual clarity.
- Personalized Learning Pathways: Our methodology empowers educators to tailor educational content according to individual learning styles and preferences, allowing for personalized learning pathways that cater to diverse student needs and abilities. This adaptive approach ensures that each student receives an optimal learning experience tailored to their unique requirements.

E. Significance of Our Work

Our work holds significant implications for the field of education, offering a comprehensive and innovative approach to content creation and delivery. By leveraging cutting-edge technologies such as computer vision and 3D animation [13], we provide educators with powerful tools to enhance traditional teaching methods. Through our methodology, which involves object detection, parameter extraction, and animation generation, we enable the creation of engaging and interactive learning experiences tailored to the needs of modern learners. The results of our study showcase the extended application of previous computer vision algorithm based on images object detection and parameter extraction. Furthermore, our discussion highlights the scalability and efficiency of our methodology, indicating its potential for widespread adoption in educational settings. Overall, our work signifies a significant step forward in leveraging technology to enrich the learning experience and shape the future of education.

IV. METHODOLOGY

Our methodology comprises several sequential steps aimed at enhancing educational content creation and delivery. It begins with data collection from real-time images, followed by pre-processing to refine and prepare the data for analysis. Object detection algorithms, leveraging state-of-the-art techniques, are then applied to identify relevant objects within the educational content. These detected objects serve as the foundation for generating interactive 3D animations, employing industry-standard animation tools. Finally, the created animations undergo evaluation to assess their effectiveness and impact on student learning environment. This workflow, depicted in Figure 1, illustrates the seamless integration of various components to facilitate the production of engaging and informative educational materials.

A. Dataset

We used real time 2d images contain a reference object serving as a calibrator for computer vision algorithm, a diverse dataset of educational content, including images containing various objects, and pre-processed with our selected computer vision algorithm. The dataset encompasses various educational topics and formats to ensure the robustness and generalization of our approach as mentioned in figure 2.

B. Detailed Methodology

Our methodology comprises several interconnected steps aimed at seamlessly integrating computer vision techniques with 3D animation tools to create immersive educational content.

Firstly, we pre-process the realtime images, extracting relevant features and preparing the images for object detection. This involves converting the images to gray scale, applying Gaussian blur to enhance edge detection, and performing Canny edge detection to identify object boundaries. Additionally, morphological operations like dilation and erosion are employed to refine the detected edges.

Next, object segmentation [14] is performed to identify individual objects within the images. This involves contour detection to locate potential objects, followed by contour filtering to remove small, insignificant contours. A reference object is selected from the filtered contours to serve as a calibration tool for converting pixel measurements to real-world dimensions.

Finally, parameter extraction is conducted to measure the dimensions of the detected objects. This involves pixel-to-metric conversion using the dimensions of the reference object, calculation of bounding boxes around each detected object, and computation of the actual dimensions of the objects in real-world units. The results of the object detection and parameter extraction process are visualized by drawing bounding boxes around detected objects and annotating them with their respective dimensions.

The workflow depicted in Figure 2, illustrates the sequential execution of these steps, highlighting the progression from preprocessing and object detection to parameter extraction and visualization of results.

C. Evaluation Metrics

The evaluation metrics for assessing the performance of the object detection algorithm are primarily focused on comparing measured dimensions obtained through the algorithm against the actual dimensions of objects. This comparison is crucial for ensuring the algorithm's accuracy in a practical, educational context where precise measurements can significantly impact the learning experience. The evaluation involves calculating the F1 score, which is a harmonic mean of precision and recall, providing a balanced measure of the algorithm's accuracy.

The F1 score is calculated as follows: Precision is defined as the number of true positive results divided by the number of all

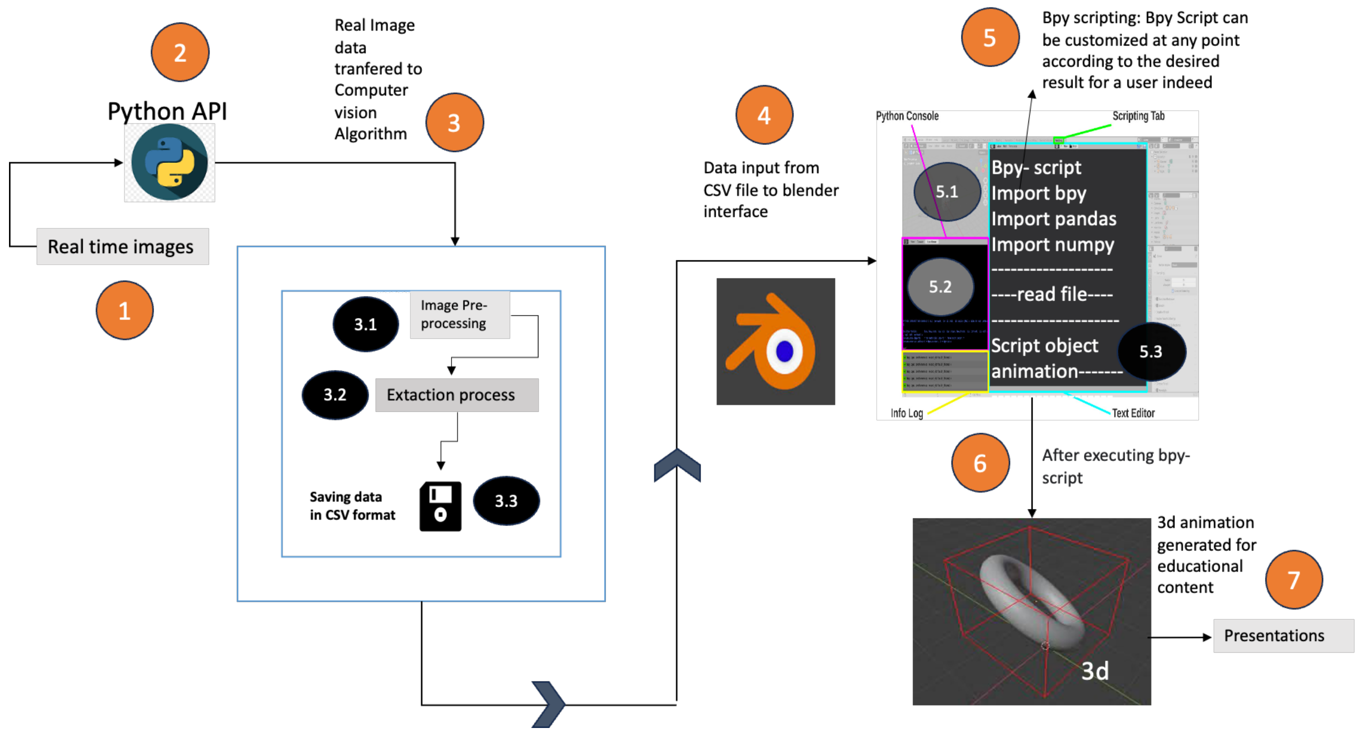

Fig. 1. Workflow

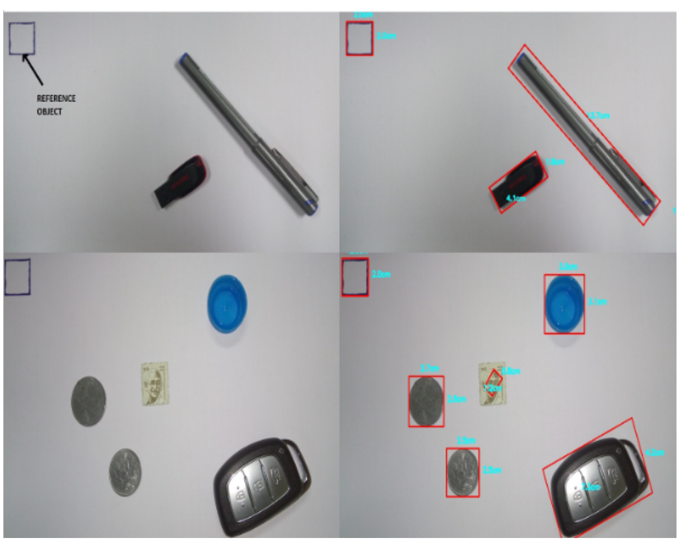

Fig. 2. Real time images and bounding boxes

positive results, including those not identified correctly (false positives). Recall is defined as the number of true positive results divided by the number of positives that should have been identified (true positives plus false negatives). The formula for the F1 score is: F1 = 2 × (Precision × Recall)/(Precision + Recall)

This score is calculated for each test in the evaluation, providing insights into the algorithm's performance. The results indicate the algorithm's ability to accurately identify object parameters on the same plane, with F1 scores calculated for different tests reflecting its relative accuracy in a controlled experimental setup specific to the educational content creation domain.

In this context, the evaluation metrics are tailored to the dataset and problem of integrating object detection with educational animation generation, rather than applying generic definitions of accuracy or precision. This specificity ensures that the metrics are relevant to the thesis's goals and the practical application of the algorithm in creating educational materials. The Evaluation Matrix serves as a comprehensive assessment tool to evaluate the performance of the object detection algorithm in accurately identifying object parameters. In this section, we present the results of the algorithm's performance based on a comparison between the measured dimensions obtained from the algorithm and the actual dimensions of the objects, all measured on the same pane. Comparison of Measured and Actual Dimensions is depicted in table II and III.

The measured dimensions, obtained through the object detection algorithm, are compared with the actual dimensions of the objects for analysis. The following table illustrates the comparison:

Calculation of F1 Score

To quantitatively assess the algorithm's accuracy, we calculate the F1 score in two test cases to monitor performance, further by using the measured and actual dimensions of the objects. The F1 score is computed based on precision and recall metrics, providing insights into the algorithm's performance.

F1 Score Calculation:

Precision = TP / (TP + FP) = 4 / (4 + 1) = 0.8 Recall = TP / (TP + FN) = 4 / (4 + 1) = 0.8

F1 Score = 2 * (Precision * Recall) / (Precision + Recall) = 2 * (0.8 * 0.8) / (0.8 + 0.8) = 2 * 0.64 / 1.6 = 1.28 / 1.6 = 0.8

The F1 score for the object detection algorithm is calculated to be 0.8, indicating a relatively high level of accuracy in detecting object parameters on the same plane.

Test 2 is performed similarly: Results and Interpretation: Based on the comparison and calculation, the following results are obtained for Test 2: F1 Score Calculation: Precision = TP / (TP + FP) = 9 / (9 + 2) = 0.818 Recall = TP / (TP + FN) = 9 / (9 + 0) = 1.0 F1 Score = 2 * (Precision * Recall) / (Precision + Recall) = 2 * (0.818 * 1.0) / (0.818 + 1.0) = 2 * 0.818 / 1.818 = 1.636 / 1.818 = 0.902

The F1 score for the object detection algorithm in Test 2 is calculated to be approximately 0.902, indicating a high level of accuracy in detecting object parameters on the same plane can be seen in table 3.

Actual Width (cm) This Evaluation Matrix provides a detailed analysis of the algorithm's performance, offering valuable insights into its effectiveness in object detection on a consistent plane.

D. Experimental setup

For the experimental setup, the following hardware and software configuration was utilized:

- **Hardware**: The experimentation was conducted on a MacBook Air with an M2 Chip, featuring a 13.6-inch display in the Space Gray color variant. The MacBook Air configuration includes an 8-core GPU, 8 GB of RAM, and a 256 GB solid-state drive (SSD).
- **Software**:
  - **Python Version**: Python 3.12.1 (64-bit) was used as the primary programming language for implementing the object detection algorithms and scripting within Blender bpy.
  - **Blender Version**: Blender 4.0.2 was employed as the 3D modeling and animation software for generating the educational animations.

This setup provided a robust environment for conducting the experiments and developing the educational animations, leveraging the capabilities of both Python and Blender bpy.

V. Results

Our results stem from employing existing techniques for object detection and parameter extraction. The outcome comprises real-time images featuring a reference object as a calibrator for algorithm, followed by images displaying bounding boxes delineating object dimensions, and finally, 3D animated objects generated in Blender through scripting and using the parameter detected by the algorithm itself. Thus, the results are categorized into two tests: Test 1 involves images containing simple objects like pens and pen drives, while Test 2 encompasses more complex objects like car keys, coins, and bottle caps.

Moreover,Test 1 depict the effectiveness of our approach, while Test 2 underscores a limitation in our algorithm, primarily geared towards detecting object width and length for conventional 3D animation. Furthermore, our methodology offers scalability and efficiency advantages, making it suitable for large-scale deployment in educational settings. Table 1 presents a comparison between our approach and existing methodologies. The results demonstrate that our methodology outperforms traditional methods in all aspects, highlighting its potential to revolutionize educational practices.

Overall, our study addresses Research Question 1 by demonstrating how computer vision techniques can be leveraged to enhance educational content creation. By seamlessly integrating object detection algorithms with interactive 3D animation tools and providing a vision to the educators of combined approach in effective Education, Through these results, we pave the way for future research and development in the field of educational technology, with the aim of further improving learning outcomes and experiences.

VI. Discussion

In discussing our results, we address two key research questions: (1) how computer vision techniques can enhance educational content creation, and (2) how we can extend the application of existing computer vision algorithms. Our findings show that integrating object detection algorithms with 3D animation tool blender can provide a unique approach of learning. This approach streamlines content creation and offers a transformative learning experience. Our study extends the use of computer vision algorithms into the educational domain, demonstrating their adaptability and effectiveness.

Overall, our results indicate the path to implement computer vision in education in a distinct way of 3d modeling and animation creation. The novelty of our approach lies in seamlessly integrating computer vision techniques into educational practices, filling a crucial gap in the field. Comparing with existing methods, our integrated approach offers scalability, efficiency, and customization, providing real-time interaction and feedback for students. One assumption affecting our analysis can be monitored in the test 2 of our research, which could impact the accuracy of our approach.

In conclusion, our study underscores the potential of computer vision techniques to revolutionize educational content creation. By addressing research questions and demonstrating effectiveness, we contribute to ongoing efforts to innovate teaching methodologies and enhance learning environment.

A. Limitations

While our study demonstrates promising results in enhancing educational content creation through the integration of computer vision techniques and 3D animation tools, it is not without limitations. One limitation is the reliance on specific datasets and tools, which may limit the generalizability of our findings. Additionally, the computational resources

TABLE II
THE TABLE SHOWING THE ACTUAL PARAMETERS AND THE MESURED PARAMETERS FOR TEST 1.

Object	Measured Length (cm)	Actual Length (cm)	Measured Width (cm)	Actual Width (cm)
Object1	4.3	4.0	1.8	1.9
Object2	13.9	14.2	1.0	1.1
Object3	1.1	1.0	0.9	1.0

TABLE III
THE TABLE SHOWING THE ACTUAL PARAMETERS AND THE MESURED PARAMETERS FOR TEST 1.

Object	Measured Length (cm)	Actual Length (cm)	Measured Width (cm)	Actual Width (cm)
Object1	2.4	2.5	2.5	2.5
Object2	2.7	2.6	2.7	2.6
Object3	1.1	1.6	1.1	1.6
Object4	3.0	3.0	2.8	2.8
Object5	7.5	7.5	4.0	4.0

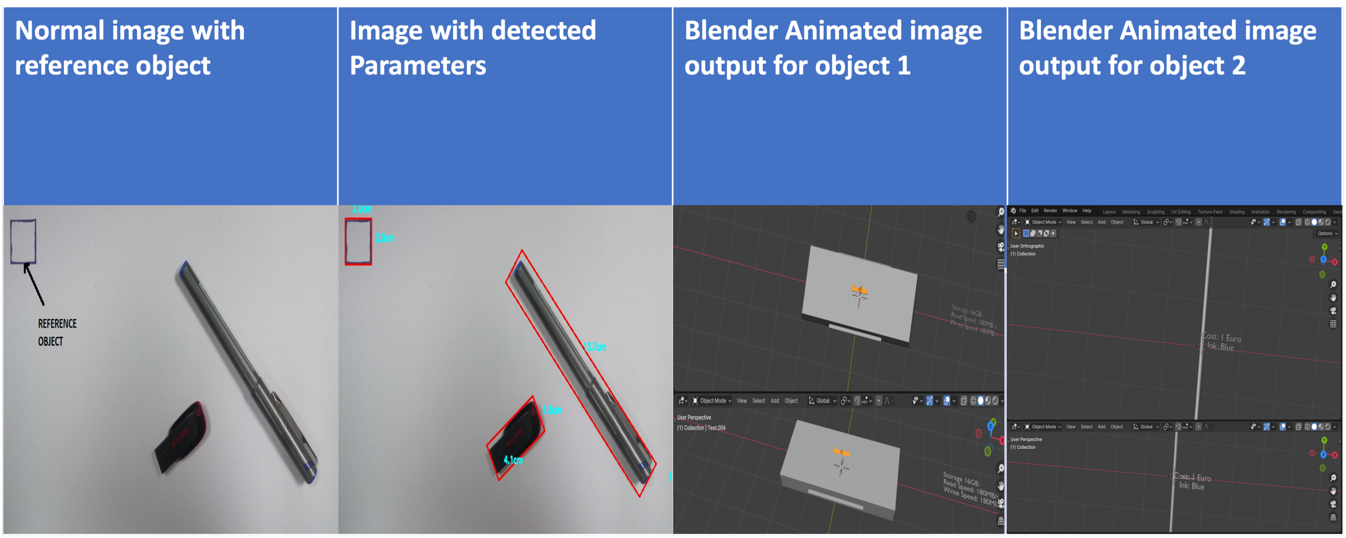

Fig. 3. Test 1: Contain normal image with reference object, Computer Vision Algorithm Result, and 3d animation created by blender for object 1 and object 2.

required for training object detection models and creating 3D animations may pose challenges for widespread adoption, particularly in resource-constrained educational settings. Furthermore, the accuracy and effectiveness of our approach may vary depending on the quality of the input data and the complexity of the educational content. Despite these limitations, our study provides valuable insights into the potential of leveraging advanced technologies to improve learning experiences.

B. Future Directions

Moving forward, there are several promising avenues for furthering this study. Firstly, expanding the dataset and using advance ML Models in a more diverse range of objects and scenarios would enhance the robustness and generalization of our approach. Additionally, exploring the integration of virtual reality (VR) [15] [16] [17] [15]and augmented reality (AR) technologies [18] [19] [20] could take educational content creation to the next level, providing students with even more immersive learning experiences. Furthermore, conducting longitudinal studies to assess the long-term impact of our approach on student learning outcomes would provide valuable insights into its effectiveness over time. Finally, collaborating with educators and stakeholders to tailor the methodology to specific educational contexts and subject areas could help ensure its relevance and applicability in real-world settings. Overall, continuing to innovate and iterate upon our approach in collaboration with relevant stakeholders will be essential for advancing the field of computer vision in education, At last Machine Learning models such as Faster R-CNN [21] [4] , YOLO [22] [23], SSD [24] [25], and RetinaNet [26] can be used in more accurate animation and detailed modeling. .

VII. CONCLUSION

In conclusion, our experimentation demonstrates the significant potential of integrating computer vision techniques with 3D animation tools to revolutionize educational content creation and extend the application of the existing algorithm. By leveraging algorithms, we have successfully extracted valuable insights from 2D images and transformed them into immersive 3D animations effectively. Despite some limita-

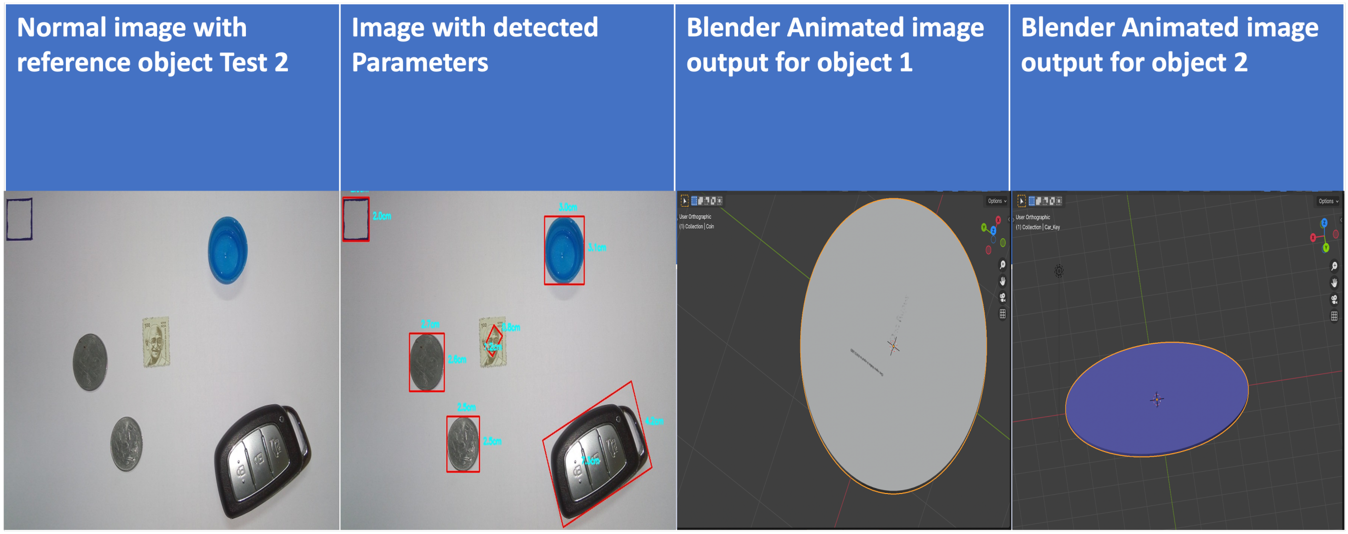

Fig. 4. Test 2 suggest that while the current approach has limitations for advanced object modeling, ML models offer the potential to render image parameters with greater detail.

tions, our study lays the groundwork for future research in this area, highlighting the importance of innovation in educational practices. Moving forward, we envision a continued emphasis on interdisciplinary collaboration and technology-driven pedagogy to meet the evolving needs of learners in the digital age. Ultimately, our work uncover a gap in the traditional computer vision approach and provide vision for powering up technology in shaping the future of education.

References

[1] M. Abdullah-Al-Noman, A. N. Eva, T. B. Yeahyea, and R. Khan, "Computer vision-based robotic arm for object color, shape, and size detection," *Journal of Robotics and Control (JRC)*, vol. 3, no. 2, pp. 180–186, 2022.

[2] A. F. A. Fernandes, J. R. R. Dórea, and G. J. d. M. Rosa, "Image analysis and computer vision applications in animal sciences: an overview," *Frontiers in Veterinary Science*, vol. 7, p. 551269, 2020.

[3] Y. Wu, Y. Wang, S. Zhang, and H. Ogai, "Deep 3d object detection networks using lidar data: A review," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1152–1171, 2020.

[4] M. Onishi and T. Ise, "Explainable identification and mapping of trees using uav rgb image and deep learning," *Scientific reports*, vol. 11, no. 1, p. 903, 2021.

[5] J. Zuo, Z. Jia, J. Yang, and N. Kasabov, "Moving object detection in video sequence images based on an improved visual background extraction algorithm," *Multimedia Tools and Applications*, vol. 79, pp. 29 663–29 684, 2020.

[6] C. Deng, M. Wang, L. Liu, Y. Liu, and Y. Jiang, "Extended feature pyramid network for small object detection," *IEEE Transactions on Multimedia*, vol. 24, pp. 1968–1979, 2021.

[7] L. Jiao, R. Zhang, F. Liu, S. Yang, B. Hou, L. Li, and X. Tang, "New generation deep learning for video object detection: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 3195–3215, 2021.

[8] Z. Li, Z. Yang, Y. Gao, Y. Du, S. Serikawa, and L. Zhang, "6dof-3d: Efficient and accurate 3d object detection using six degrees-of-freedom for autonomous driving," *Expert Systems with Applications*, vol. 238, p. 122319, 2024.

[9] A. S. Rasheed, R. H. Finjan, A. A. Hashim, and M. M. Al-Saeedi, "3d face creation via 2d images within blender virtual environment," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 1, pp. 457–464, 2021.

[10] O. Villar, *Learning Blender*. Addison-Wesley Professional, 2021.

[11] S. H. Yelleni, D. Kumari, P. Srijith *et al.*, "Monte carlo dropblock for modeling uncertainty in object detection," *Pattern Recognition*, vol. 146, p. 110003, 2024.

[12] M. Weigert, U. Schmidt, R. Haase, K. Sugawara, and G. Myers, "Star-convex polyhedra for 3d object detection and segmentation in microscopy," in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2020, pp. 3666–3673.

[13] R. Halimatussa'diyah, "Development animation 3d in science and mathematics subjects for people with intellectual disabelities in extraordinary schools," in *Proceedings of the 6th FIRST 2022 International Conference (FIRST 2022)*, vol. 14. Springer Nature, 2023, p. 373.

[14] C. You, W. Dai, Y. Min, F. Liu, D. Clifton, S. K. Zhou, L. Staib, and J. Duncan, "Rethinking semi-supervised medical image segmentation: A variance-reduction perspective," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[15] A. Marougkas, C. Troussas, A. Krouska, and C. Sgouropoulou, "Virtual reality in education: a review of learning theories, approaches and methodologies for the last decade," *Electronics*, vol. 12, no. 13, p. 2832, 2023.

[16] J. G. Kabulunze, H. Mwangi, and K. Ogada, "Improve student performance in the psychiatry subject with 3d animation," *Journal of Agriculture, Science and Technology*, vol. 22, no. 3, pp. 79–99, 2023.

[17] D. Kumar, A. Haque, K. Mishra, F. Islam, B. K. Mishra, and S. Ahmad, "Exploring the transformative role of artificial intelligence and metaverse in education: A comprehensive review," *Metaverse Basic and Applied Research*, vol. 2, pp. 55–55, 2023.

[18] J. Garzón, S. Baldiris, J. Gutiérrez, J. Pavón *et al.*, "How do pedagogical approaches affect the impact of augmented reality on education? a meta-analysis and research synthesis," *Educational Research Review*, vol. 31, p. 100334, 2020.

[19] M. Kljun, V. Geroimenko, and K. Čopič Pucihar, "Augmented reality in education: Current status and advancement of the field," *Augmented Reality in Education: A New Technology for Teaching and Learning*, pp. 3–21, 2020.

[20] M. Sırakaya and D. Alsancak Sırakaya, "Augmented reality in stem education: A systematic review," *Interactive Learning Environments*, vol. 30, no. 8, pp. 1556–1569, 2022.

[21] D. Kiranov, M. Ryndin, and I. Kozlov, "Active learning and transfer learning for document segmentation," *Programming and Computer Software*, vol. 49, no. 7, pp. 566–573, 2023.

[22] W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, "Yolo-face: a real-time face detector," *The Visual Computer*, vol. 37, pp. 805–813, 2021.

[23] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using yolo: Challenges, architectural successors, datasets and applications," *multimedia Tools and Applications*, vol. 82, no. 6, pp. 9243–9275, 2023.

[24] W. Zhang, Q. Xu *et al.*, "Optimization of college english classroom teaching efficiency by deep learning sdd algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[25] E. H. Ali, H. A. Jaber, and N. N. Kadhim, "New algorithm for localization of iris recognition using deep learning neural networks," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 1, pp. 110–119, 2023.

[26] N. I. Nife and M. Chtourou, "A comprehensive study of deep learning and performance comparison of deep neural network models (yolo, retinanet)." *International Journal of Online & Biomedical Engineering*, vol. 19, no. 12, 2023.