Review

Computer Vision Meets Educational Robotics

Aphrodite Sophokleous [1][iD], **Panayiotis Christodoulou** [1][iD], **Lefteris Doitsidis** [2][iD] and **Savvas A. Chatzichristofis** [1,*][iD]

1   Intelligent Systems Laboratory, Department of Computer Science, Neapolis University Pafos,
    8042 Pafos, Cyprus; a.sofokleous@nup.ac.cy (A.S.); panayiotis.christodoulou@nup.ac.cy (P.C.)
2   Intelligent Systems & Robotics Laboratory, School of Production Engineering & Management,
    Technical University of Crete, 73100 Chania, Greece; ldoitsidis@dpem.tuc.gr
*   Correspondence: s.chatzichristofis@nup.ac.cy

Abstract: Educational robotics has gained a lot of attention in the past few years in K-12 education. Prior studies have shown enough shreds of evidence and highlight the benefits of educational robotics as being effective in providing impactful learning experiences. At the same time, today, the scientific subject of computer vision seems to dominate the field of robotics, leading to new and innovative ideas, solutions, and products. Several articles from the recent literature demonstrate how computer vision has also improved the general educational process. However, still, the number of articles that connect computer vision with educational robotics remains limited. This article aims to present a systematic mapping review, with three research questions, investigating the current status of educational robotics, focusing on the synergies and interdependencies with the field of computer vision. The systematic review outlines the research questions, presents the literature synthesis, and discusses findings across themes. More precisely, this study attempts to answer key questions related to the role, effectiveness and applicability of computer vision in educational robotics. After a detailed analysis, this paper focuses on a set of key articles. It analyzes the research methodology, the effectiveness and applicability of computer vision, the robot platform used, the related cost, the education level, and the educational area explored. Finally, the results observed are referred to as educational process benefits. The reviewed articles suggest that computer vision contributes to educational robotics learning outcomes enhancing the learning procedure. To the best of our knowledge, this is the first systematic approach that revises the educational robotics domain by considering computer vision as a key element.

Keywords: educational robotics; computer vision; educational tool

1. Introduction

In recent decades, education has been transformed and transitioned beyond the traditional learning process methods and is now enriched with procedures that make use of technological, mainly Information Communication Technology (I.C.T.), related tools. Many researchers studied the convergence of I.C.T. in education while highlighting the growing and successful incorporation between I.C.T. applications and teaching [1]. They provided clear explanations about the significance of its I.C.T. role, identifying the opportunities offered to teachers and students [2], resulting in a more useful and exciting learning process.

Over the years, rapid growth in robotics has been reported, improving a lot of developments in many fields, such as navigation and path planning [3,4], search and rescue applications [5], industrial applications [6], and entertainment. Considering the impact of the field, robots would inevitably be adapted for educational purposes also. Educational robotics is a field of study that aims to improve the learning experience through the creation, implementation, improvement, and validation of pedagogical activities [7–9]. Learning theory principles, constructivism and constructionism [10], are particularly bearing for the field of educational exploitation of robotics. According to Piaget, learning results from interaction with the environment lead to new learning experiences [7,11]. In [12], another

study is presented. Authors report the benefits of the Internet of things (IoT) adoption in education, including increased interactivity, personalized learning, efficient classroom management, and better student monitoring. In [13], the integration of computer vision in tandem with IoT in education is also discussed.

The importance of Science, Technology Engineering, and Mathematics (STEM) and, therefore, educational robotics as a core part of it has been identified as a key tool towards digitizing education, involving students in learning activities, and developing their skills project-based learning through robotics. In this context, in recent years, many research programs have emerged related to educational robotics. Among them, the 'Educational Robotics for ER4STEM (STEM)' which has a concept the three important pillars of constructionism: (i) engaging students with powerful ideas, (ii) building on personal interests, and (iii) learning through making (or presenting ideas with tangible artifacts). Edubots is an ongoing Erasmus+ Knowledge Alliance project aiming to improve results and raise attainment levels in European higher education. The 'Science with robotics' project is another Erasmus+ Knowledge Alliance project which provides teachers with the necessary training to go one step further and introduce robotics to work with content from different areas. The CODESKILLS4ROBOTICS project, also funded by Erasmus+, aims to design an innovative program that aims to introduce coding and robotics' to primary school students. The direct target group is children aged 9–12 years old, emphasizing children with fewer opportunities who will learn how to code. Furthermore, of course, these are just a small indicative sample of the research projects that have been funded and helped the evolution, adoption, and spread of educational robotics.

The increased importance of educational robotics for the academic and educational communities has been identified in [14]. The authors in [15] define the educational robotics term according to the three main fields involved: (1) education, (2) robotics, and (3) human–computer interaction and conclude that educational robotics '. . . is a field of study that aims to improve the learning experience of people through the creation, implementation, improvement, and validation of pedagogical activities, tools (e.g., guidelines and templates) and technologies, where robots play an active role, and pedagogical methods inform each decision. . .'.

The authors in [7] reassess the definition of educational robotics, using a bibliometric map, as an '. . . an essential branch of educational technology implemented by activities designed using the theory of constructionism focused on the development of computational thinking skills, collaborative learning, and project-based learning. . .'. Various definitions are also presented in [16,17]. In [16], educational robotics is described as a '. . . research field that aims to promote active engage learning through the artifacts students, create the phenomena and simulate. . .' while, in [17] educational robotics is defined as '. . .the application of robots and robotics activities in teaching and learning. . .'.

According to the literature, there are various ways where educational robotics can be applied in the learning process: (i) as a learning and teaching tool, during the pedagogical method or an educational practice where robotics are used as another I.C.T. tool in the hands of teachers; (ii) as a cognitive, educational object, where robotics is just another subject with its curriculum, and the student learns and understands among others the concept of robotics, the technical knowledge on how it works, how it is programmed and how it can be managed; (iii) as social robots where they interact naturally with humans and behave in a way that is comfortable for humans [15] and finally; (iv) as a valuable tool that can help students in developing cognitive and social skills during their K-12 education [18]. Independent of how educational robotics is integrated into the learning process, they aim to fulfill certain learning outcomes, as formally defined and outlined in [7]. These help to:

- Improve problem-solving skills by helping the student understand difficult concepts more easily, research, and conduct decisions.
- Increase self-efficacy: The machine's natural handling promotes experimentation, discovery, and rejection and, consequently, enhances the student's self-confidence

because the student feels that he controls the machine. This also strengthens the students' critical thinking.

- Improve computational thinking: Students acquire algorithmic thinking to break down a large problem into smaller ones and then solve it. Students learn how to focus on important information and reject irrelevant ones.
- Increase creativity by learning with play-transmitting knowledge in a more playful form. Learning turns into a fun activity and becomes more attractive and interesting for the student.
- Increase motivation as educational robotics enables students to engage and persist at a particular activity.
- Improve collaboration as the team spirit and the cooperation between the students are promoted.

Currently, there is a broad range of robots for serving different requirements and age groups among students [19] ready to be used during the educational process. Although educational robotics and computer vision, as different fields, may be part of the educational process in K-12 education until recently, no one has ever investigated how they jointly may support the educational robotics area. This concern has risen since computer vision, one of the key tools used by the research community in robotics has significantly contributed to adopting various robots in different applications and introducing them as mainstream devices.

To fully understand the importance of computer vision in education and specifically in educational robotics, it is deemed necessary to define computer vision. According to the literature, there are many definitions of computer vision. In [20], it is defined as the '... science that studies means on how to provide to a computer the ability to 'see'...' computer vision uses cameras to analyze or understand scenes in the real world [21] and allows computers to capture, interpret, process the visually perceivable objects, and understand the captured digital images and react suitably. Moreover, during recording light on a video camera, computer vision can be defined as the scientific field that extracts information from digital images that can eventually lead to a decision or execution of an action. Computer vision deals with how computers can acquire high-level knowledge from digital content.

Nowadays, the number of personal, medical, scientific, and social networking images uploaded on the Internet is growing exponentially. Computer vision is essential because we need computers to understand the images' content, describe the real-world that humans see in one or more images, and reconstruct its properties, such as shape, illumination, and color distributions [22]. Moreover, as distance learning and online classrooms require good quality of both image and streaming video, recent advances in computer vision and algorithms have made considerable potential improvements [20]. An endless list of fast-growing and advanced computer vision applications is being used today in a wide variety of real-world applications, including sports, health and medicine, agriculture and farming, autonomous driving, social distance, people counting, and so on.

One of the essential educational principles for both educators and students is how knowledge is constructed. According to the sociology of education based on the individual's uniqueness, everyone learns differently [23]. The authors in [24] present computer vision to improve learning and knowledge acquisition. Teaching methods can be enhanced through computer vision tasks by analyzing the students' interest level, body posture, eye movement, and behavior. Subsequently, teachers can immediately react by modifying their teaching methods to harvest more attention from students, maximize their interest, and design lectures that are easier to understand [13]. In addition, computer vision in education can maximize students' academic output by offering customized learning experiences based on students' strengths and weaknesses. Moreover, it can improve students' and teachers' relationships, especially for students with learning difficulties. Between 2014 and 2020, 111,100 articles with the keyword 'computer vision in education' have been published. Figure 1 depicts the upward trend of research in this field.
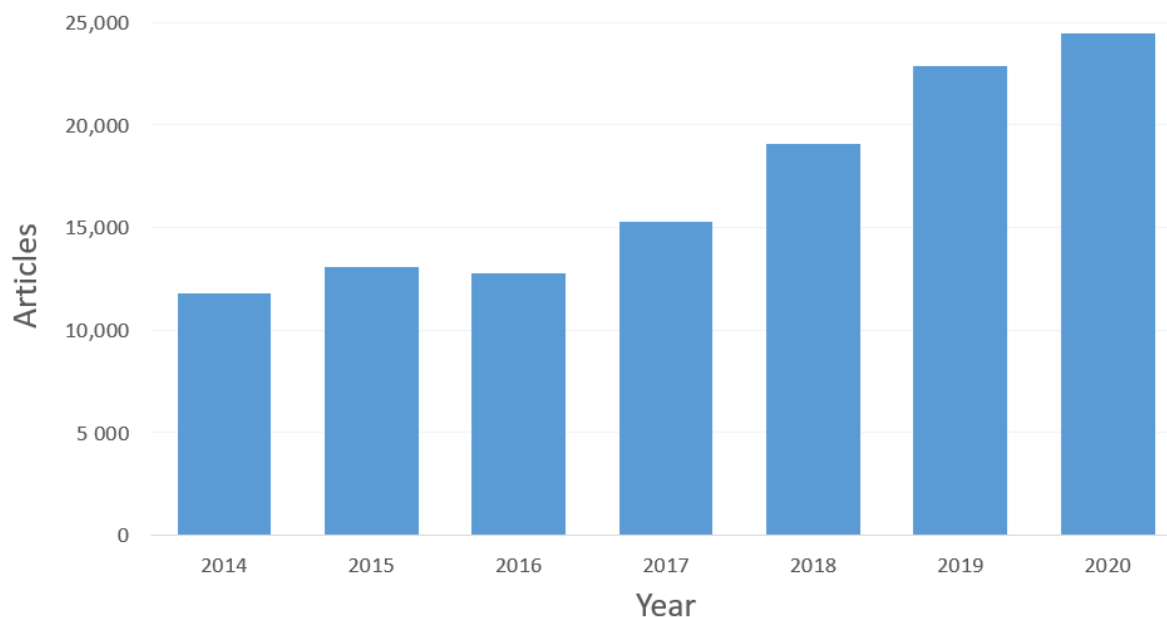
Figure 1. Articles per year reporting in computer vision in education.

Motivation

Despite the recent research attention on applying computer vision in education, there is still a limited amount of works focusing on educational robotics' applications. This paper revisits the articles that adopt computer vision mechanisms and technologies on educational robotic tasks, highlighting the impact of combining these two scientific fields. This work's primary focus is, to provide a systematic review, shaping an overview of computer vision and educational robotics' current research. The article aims to investigate how computer vision enhances and supports the educational robotics' impact. It examines the role of computer vision in educational robotics and the benefits of using computer vision in K-12 education. Moreover, it checks how easy and affordable it is to integrate computer vision and educational robotics in K-12 education. Finally, this study identifies and determines computer vision's role in the learning procedure and how it improves students' interest and performance in K-12 education. Overall, this study attempts to answer key questions related to the role, effectiveness and applicability of computer vision in educational robotics.

The rest of the paper is organized as follows. Section 2 firstly outlines the adopted research methodology and the process of collecting relevant research papers and then presents the gathered papers' results. Section 3 analyses and summarizes the study's outcomes, and provides answers to the research questions examined in this work. Section 4 concludes the article.

2. Research Methodology

A systematic mapping study suggested by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [25] was selected as the research methodology for this study. As depicted in Figure 2, the systematic mapping procedure aims to provide an overview of a research area, identify if research evidence exists, and quantify the amount of evidence. In order to accomplish our goal we follow the systematic review process described by [8,9,26,27]. The systematic review outcomes will help us identify and map research areas related to computer vision and educational robotics and possible research gaps.
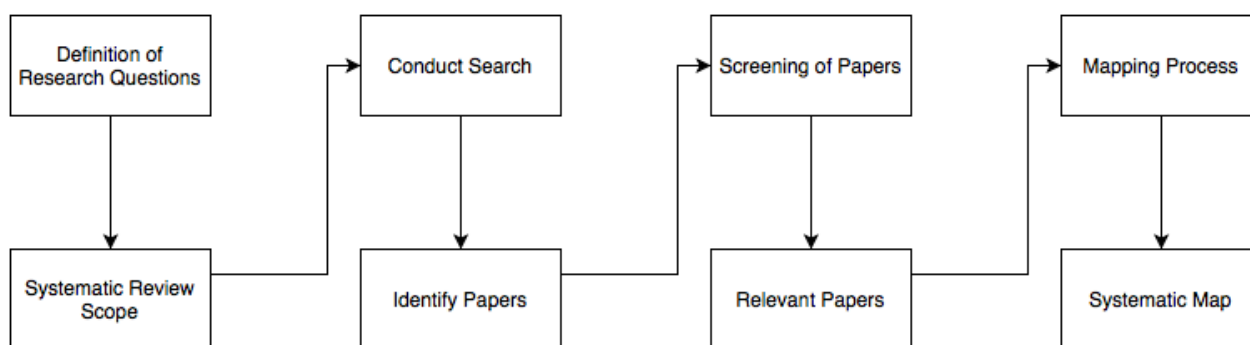
Figure 2. The systematic mapping process.

2.1. Definition of Research Questions

The first stage of the systematic mapping process is the definition of the research questions. As the primary goal of this systematic review is to identify the synergies and interdependence between computer vision and educational robotics, we developed a bibliometric map that helped us define this study's research questions.

The bibliometric map, presented in Figure 3, was constructed considering two criteria as reference points, computer vision and educational robotics. A set of related keywords to the critical criteria are represented on a 2-Dimensional plane. On a bibliometric map, the keywords that co-occur are linked together through a line that has a length proportional to the co-occurrences; this outlines the similarity (link strength) between the terms. The distances that may exist between the various keywords on the map compromise indicators of dissimilarity. As it can be observed from Figure 3, the map constructs a triangle of three main classes: (1) computer vision, (2) robotics, and (3) educational robots. Each of the classes mentioned above creates various keywords which are analyzed further below.

Near the computer vision class, we can observe several other keywords, including robots, object recognition, image processing, speech recognition, and other visual serving terms that mainly compose the definition of computer vision. The educational robots class is related to deep learning, convolutional neural networks, learning systems, virtual reality, deep neural networks, artificial intelligence, intelligent robot, and robot learning. The conjunction between computer vision and educational robots' classes raises questions about how computer vision is linked with educational robots and how it can aid the overall educational process. The third broader class of robotics, consists of the following keywords curricula, cost, object detection, and cameras.

Another aspect that can be observed from the bibliometric map analysis is that near the educational robotics class which is the result of the robotics and educational robots classes, there are additional keywords like teaching, students, robot programming, image processing, object recognition, and education. This raises new questions concerning (1) How educational robots use image processing or object recognition to enhance the teaching process, and (2) What benefits do students receive from educational robots.

Finally, it is worth mentioning that on the one hand, social robots as well as humanoid robots are close to the educational robots and educational robotics classes, respectively. On the other hand, the keywords human–computer interaction and human–robot interaction are close to the computer vision class. This observation concludes that most robotics platforms that adapt computer vision mechanisms are humanoid social models. This observation was also taken into consideration during the formation of the research questions below.
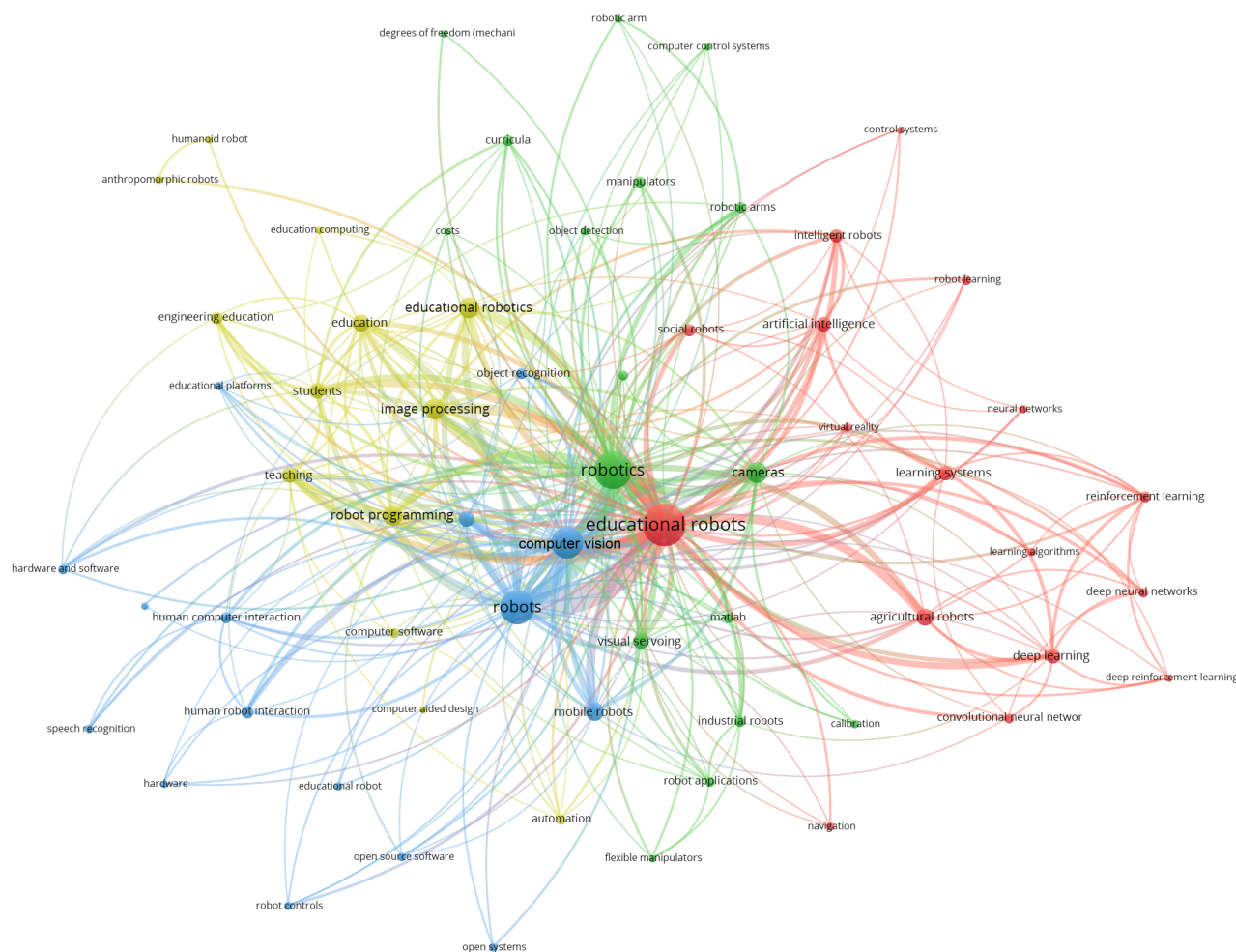
Figure 3. Overview of *Q*1 index term bibliometric map based on Scopus data.

The observations above and questions were used to set up the main research questions which guided us through the systematic review. These are the following:

- RQ1: What is the role of computer vision in educational robotics?
  The first research question helps the reader identify the current research that has been conducted on computer vision in educational robotics and attempts to provide answers on how computer vision can be used during the learning process.
- RQ2: How computer vision benefits educational robotics' expected learning outcomes in K-12 education?
  The second research question revisits educational robotics' expected learning outcomes and investigates how computer vision benefits K-12 education.
- RQ3: How affordable and feasible is the integration of solutions that combine educational robotics and computer vision in K-12 instructional activities?
  The third research question aims to reveal if the integration of computer vision and educational robotics activities in education can be adopted from a cost–benefit perspective. The same research question investigates also the ease of access and the availability of tools.

2.2. Search Approach

A detailed search protocol was established to identify all scientific papers of interest for our study. Our goal was to reduce, if not eliminate, the possibility of researcher bias. Before finalizing the appropriate search keywords for this study, we conducted pilot searches and tested possible keywords. We concluded in using the following query Q_1 as the search terms:

$Q_1=$

('image processing' **OR** 'camera' **OR** 'computer vision')

AND

('robotics in education' **OR** 'educational robotics' **OR** 'educational robots' **OR** 'robotics learning' **OR** 'robotics teaching')

We extracted high-quality peer-reviewed papers published in various conferences and journals related to the research topic. For paper retrieval, we used the following scientific databases (1) IEEE Xplore, (2) ACM Digital Library, (3) Springer Link, and (4) ScienceDirect.

2.3. Screening of Relevant Papers

Utilizing the Q_1 query, we retrieved 370 related papers. The yearly distribution of the papers, from 2014 to 2020, is presented in Figure 4. Given that many of those papers were not implicitly related to the research questions, we needed to assess them for actual relevance.
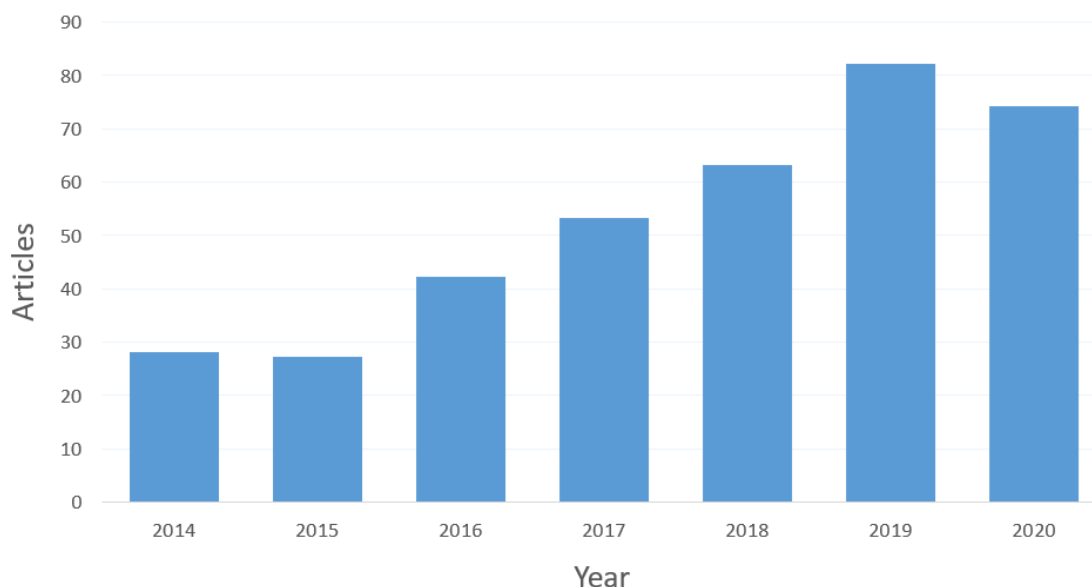

Figure 4. Articles per year reporting in computer vision in educational robotics.

Since our primary focus is on papers that use computer vision in educational robotics from pre-schools to secondary education, we identified a set of inclusion and exclusion criteria to shortlist the papers used to answer our research questions.

Inclusion Criteria (I.C.):

- I.C.1: Articles that present the use of computer vision in schools along with experimental outcomes.
- I.C.2: Articles that outline computer vision as an educational tool from pre-schools to the high school context.
- I.C.3: Articles that present computer vision as an assistive tool to support the educational process from pre-school to the high school context.

Exclusion Criteria (E.C.):

- E.C.1: Articles that did not mention the use of computer vision in educational robotics.

- E.C.2: Articles that outline the positive effects of computer vision in education, but they do not provide experimental results.
- E.C.3: Articles that are related to higher education.
- E.C.4: Articles that mention the use of robotics in education without utilizing computer vision techniques.
- E.C.5: Theses or books or annual reports.
- E.C.6: Articles that describe teachers' efforts on educational robotics.
- E.C.7: Articles that were not written in English.
- E.C.8: Articles published before 2014.

2.4. Mapping Process

Based on the Inclusion and Exclusion Criteria defined during the screening of relevant papers, our initial set of 370 articles was reduced to 21 articles. In total, 176 articles were excluded because they were focusing on computer vision tasks (object recognition, image segmentation, color recognition, and convolutional neural networks) in other scientific fields such as Engineering, Medicine, etc., and not in educational robotics. Furthermore, 50 of the articles were removed as they used robotics in education but did not utilize any computer vision techniques. Subsequently, 39 articles were excluded as they do not show any experimental results to support their study and 31 papers were removed since they describe computer vision or education robotics in higher education. Apart from that, 30 articles were excluded because they were published as thesis, books, or annual reports and 4 papers were removed as they outlined teachers' efforts in robotics education. Finally, 9 were excluded as they were not written in English.

To sum up, the final list consists of 21 papers published from 2014 until 2020 which present the use of computer vision as an educational or assistive tool that supports the educational process from pre-school to the secondary school context (K-12 education).

The mapping process stage is divided into two steps. During the first step, we read the abstract and identified keywords that reflected the paper's contribution. During the second step, we firstly developed a higher level of understanding based on the identified keywords. We used those keywords to form the various categories, and finally, we read the selected papers. We were continuously updating the categories or creating new ones if an article was revealing something new. This process resulted in forming a systematic map of clusters that took into consideration all relevant papers.

3. Analyzing the Literature

This section analyses the systematic review results and answers the three research questions identified and highlighted during the systematic mapping process.

3.1. What Is the Role of Computer Vision in Educational Robotics?

The first research question helps us identify the current research that has been conducted on computer vision in educational robotics and tries to provide answers on how computer vision can be used during the learning process.

The authors in [28] present the use of a robotic educational system that exploits advanced computer vision capabilities to detect written characters. The histogram of oriented gradients (also known as HOG) is used as a low-lever descriptor of the characters' detection stage. The proposed system aims to help new alphabet learners, mainly young children, write alphabet characters correctly. The system was benefiting from advanced computer vision algorithms to detect written characters. While interacting with the robot, children are led to a point where they want to write clear enough to make the robot understand their handwriting or write fast to meet the robot's requirements.

In a more sophisticated setup, Wu et al. in [29] introduced a robotic educational system combined with an object recognition technology that provides innovative second language learning services for pre-school children in China. The kid places physical objects into recognizable areas for an interactive operation. An avatar guides them in English to

touch, drag, click, and press to interact with various objects. The presented system consists of three main components: a projector that casts images and items on the flat surface with which the child interacts, a Kinect that takes pictures of things on a fixed area to realize object recognition and finger tracking, and the main controller that receives the captured object of the camera and identifies and controls the content of the projector playback. The object recognition uses the SURF algorithm provided by OpenCV to obtain SURF features from the database.

Subsequently, authors in [30], demonstrated a prototype for a robotic language tutor that uses various computer vision techniques for behavior analysis, face recognition methods for guessing the user's age, object and speech recognition modules, and synthesis tools to emulate a human-to-human interaction. For this purpose, the authors used the state-of-the-art architecture GoogLeNet for object recognition and deep-convolutional neural networks for classifying age and gender trained using the Caffe framework [31]. The teaching process is adjusted according to the user's age estimation. Initially, communication between the robot and the user starts, and then object detection is being used to enable the communication of the object's name in the user's language. It is worth noting that this article does not involve any tangible device, but it was chosen to be included in the analysis since the authors classify it as a robotic educational tutor.

Moreover, Kusumota et al. [32], adapts a Cozmo mobile robot for educational purposes. The Cozmo robot utilizes computer vision using the Google Cloud Vision API. The robot receives images through web requests and returns a set of textual image characteristics. The robot was developed to run educational functions and games that include mathematical operations, spelling, directions, and question functions. In the proposed paper, various procedures for educational purposes were implemented on a web server for a more friendly user interface. More specifically, the first function tested with the students' was the drawing shapes function and Cozmo was programmed to draw a circle and a square shape. The second function tried was the sum function. When a student was finding the right answer, Cozmo played a happy animation. The last function tested was spelling. Students' had to spell their names, and when they were making mistakes, Cozmo was playing the sad animation.

The educational benefits of computer vision in educational robotics are also analyzed in [33]. In [33], the authors introduce the MonitoRE system to create an interactive educational environment for teaching robotic. MonitoRE helps students during the teaching-learning process through the web camera by conducting computer vision tasks. The image processing is done with the support of the OpenCV Computational Vision library and is used to complete different activities, like rescue activity, divided into multiple degrees of difficulty. An object is placed in a predetermined location, and then it needs to be rescued by the robot using color and object recognition.

In addition, an educational robotic system for preschoolers' cognition education based on the NAO platform is presented in [34]. The robot's model uses a fast object recognition mechanism which utilizes region proposal networks [35] and convolutional neural networks. The robot's core aims to automatically generate visual questions and answers based on the recognition results, including pronunciation, spelling, story, learning cards, and other related resources to serve as a learning trainer and partner. More specifically, objects in the real-world are detected, and a set of learning materials associated with the objects is presented to the learners. For example, when a cat is detected, the robot will teach learners to pronounce the word 'cat' in different languages, and more related pictures will be presented to the learners. For geometrical thinking training, an automatic questioning-and-answering section that implements voice interaction between learners and robots is employed to engage learners' thinking.

As observed in the articles' analysis, computer vision in educational robotics is also applied in special education. In [36], the authors introduce the use of humanoid robots, such as NAO, in special education, with emphasis on children diagnosed with autism spectrum conditions. The robots' primary goal is to encourage and improve imitation

and social-communication skills of the child by taking advantage of computer vision algorithms' capabilities. Thus, the NAO's visual system employs a localized version of the color, and edge directivity descriptor [37] and a bag of visual words model in its recognition tasks to implement simple imitation games for the therapist's objectives. Moreover, Amanatiadis et al. in [38], extended the previous studies in special education using humanoid robots by adding multi-robot game sessions therapies with two children. Computer vision algorithms based on color features and the robot operating system for inter-process and multi-robot communication were used. Children face two NAOs, the NAO1 that demonstrates a game like 'Rock-Paper-Scissors' and NAO2 which asks the children to participate. Additionally, a humanoid NAO Robot outlined by the authors in [39] was also used in therapy sessions from children with Down Syndrome. In order to deal with image processing purposes, OpenCV, a computerized visionary library, was used by the authors. In [39], the robot's purpose was to teach children how to recognize various colors using the camera by mentioning a toy's color every time they showed a humanoid NAO figure. Since then, efficient tools such as tactile and precision sensors, cameras, microphones, and voice synthesizers were used to take advantage of the capabilities of NAO because humanoid robots can attract children's attention.

A recent study demonstrates that children's interaction and communication are enhanced through computer vision mechanisms. The authors in [40] proposed, once again, using a NAO robot, a platform for teaching geometric figures and colors to children in nursery age. Children hear various color or shape names and touch the different colors or shapes on the board during the activities. Then the NAO, using computer vision mechanisms, checks what the children have chosen and either correct or reward them. In addition, another work that uses open-source robotics to support the synergistic learning of computational thinking and STEM, with an emphasis on computer science, is introduced in [41]. In the proposed work, students used a robotics learning platform that combined the physical and algorithmic aspects of model building and problem-solving through computer vision algorithms' for shape and color detection, object tracking, or face detection.

Moreover, the study presented in [42] outlines a social NAO robot that interacts with a child while its playing until the child becomes 'Happy'. The NAO robot includes a fuzzy rule-based system and sensor signals processed by Computational Intelligence and Machine Learning algorithms. In [42], the authors proposed a feedback control that compares a resultant sentence caused by crowd-computing techniques to a computer-vision-induced sentence in driving a linguistic controller. Sentences have to be correct such as 'Give the toy to an older child,' 'Give the toy to a child of the opposite gender,' 'Change Toy', etc. The 'Happy' feeling is succeeded while the child plays a game with the robot to recognize their age, expressions, and gender by conducting computer vision tasks.

The ChildBot outlined in [43] presents a different study that uses multiple robots' platforms for educational purposes. ChildBot includes several modules such as audio-visual active speaker localization, object tracking, visual activity recognition, and distant speech recognition. The integrated visual system classifies the encoded features that result from Vector of Locally Aggregated Descriptors (VLAD) by employing linear support vector machines and perceives various events during the interaction, such as children's speech and activities, children's locations in the room, and tracking of objects, and asks them to complete different tasks-games. For example, the robot requests a child to perform a gesture that usually denotes a meaning and then asks the child to confirm the recognition. Another task is the Pantomime; the child can use their whole body to mimic an activity and interact extensively with the robot. Both the robot and the child repeatedly swap the mime's roles. After a child's reaction, the robot also expresses the same feeling using its body and face.

All the aforementioned studies adopt computer vision mechanisms to enhance the educational robotics-based learning procedure strictly. Other studies engage computer vision as assistive technology to stimulate the students' interest. Of course, several approaches

combine the twofold nature of computer vision in educational robotics resulting in efficient solutions that significantly enhance the educational process.

In [44], 38 children in the ages of 10–11 were separated into two groups to solve various mathematical concepts taught (arithmetic). Group 1 performed the teacher's activities, and Group 2 performed the activities with the teacher and a robot teaching-assistant. By comparing the results of both groups, Group 2 scored better than Group 1 in all questions. In addition, experiments outlined that even not all children like mathematics, when they were learning mathematics with a robot's help, they enjoyed the lesson. Finally, most of the children in Group 2 believe that the NAO robot helped them understand the course more easily, and they all stated that they would like to have a robot-assistant in their classroom. In the previous paper, the robot's computer vision acted as a mediator to co-teach and aided the education process. The authors in [11] introduce multimodal NAO robots for learning purposes in the classroom when teaching various courses such as Danish, English, ethics, programming, and technology. Pupils mainly used the robot's text-to-speech and gesture features. The use of such robots benefits pupils' experience in both academic and technological teaching. In [11], the NAO's camera helped assistive tasks, but not in teaching.

Furthermore, the authors in [45] presented a study with 46 Iranian female students in the age of 12, who study junior English, divided into two groups. The main objective of the study was to investigate vocabulary learning through interaction with a human teacher assisted by a humanoid robot. The first group consists of 30 students who use an intelligent robotics-assisted language learning tool, known as RALL. The second group contains 16 students who do not have access to the RALL system. The RALL system consists of a NAO robot with voice command/recognition and computer vision capability, providing an opportunity for discussion, and prompt students to think of the word or concept. The paper concludes that the RALL group achieved higher scores on both the post-test and the delayed post-test.

In [46], the authors present the capability of young students to interact and communicate with a NAO robot in an autonomous way and in a teleoperated way (when someone controls the robot). Communication exists in three ways: speech, vision, and gesture. For the visual module, a combination of techniques to detect and recognize the chosen objects was selected. More precisely, the VOCUS2 system for segmentation and background noise extraction was used; then, SURF features extraction and the bag-of-words method were utilized and, finally, trained with multiple Support Vector Machines. Experiments were performed randomly by assigning 82 students aged between 7 and 11 to interact with the robot.

Educational ROS Robot Platform (EUROPA) [47], is an open-source robotic platform focused on STEM teaching that can be applied in physics, mathematics, and computer science courses. EUROPA's hardware consists of a Raspberry Pi3 B+ while its software infrastructure is based on the Robot Operating System (ROS) that covers a range of applications, from basic educational robotics to advanced applications, such as vision and mapping. Vision is performed by extracting color features from the camera's images using the OpenCV library. The vision's goal is to use the camera as a color sensor and direct EUROPA to follow a yellow line painted on the floor. In addition, through video activities, children can teleoperate the robot from their computer keyboard. Projects like the EUROPA aim to provide students with real-world STEM examples and a better understanding of notions that they have already been taught.

Subsequently, the authors in [48] employ a Bee-bot robot in pre-school education to provide immediate, personalized feedback and recommendations to young children while performing a series of programming-related activities. The proposed system uses an intelligent fuzzy-rule-based system and computer vision techniques to monitor the activities and interact with the participants. These activities are related to algorithmic thinking and sequencing. Participants were divided into three groups: the first one that used a computer graphical interface, the second which provided the instruction directly

to the tangible robot, and the third group that adopted a hybrid approach composed of the Bee-bot and the proposed computer vision platform. Participants were receiving instructions through simple stories such as . . . 'robot has to go to school starting from her house assist it by giving her the correct instructions to take the shortest path and not be late'. . . or . . . 'After school, the robot must visit the grandmother's house to have lunch with her. However, they must be cautious, avoid the factories as they are hazardous places for a young robot', and interacting with the robot.

The contribution of computer vision in teaching STEM is also highlighted in the following articles. The PiBot project, described in [49], was developed to improve robotics' teaching in secondary education. PiBot is an open low-cost robotic platform with computer vision capabilities used in the classroom to train pre-university students during STEM education. Image processing is performed using the OpenCV library, a standard in the computer vision community. The activities of PiBot cover programming, robotics, and technology.

The work outlined by the authors in [50] describes a robot platform that aims to help the student learn how to code using a more exciting methodology that makes the student more interactive in solving problems. The robot uses a single camera complemented with the following computer vision algorithms (a) a field detection algorithm, (b) a robot position and orientation detection, and (c) a robot neighborhood extraction and labeling algorithm. Students deal with algorithms that help the robot to prevent obstacles to reach the goal point.

In summary, based on the literature, computer vision can enhance educational robotics activities and learning procedures, following two different learning mechanisms. On the one hand, several papers proposed specially designed computer vision tasks to enhance the learning process. In this category, computer vision undertakes, for example, to recognize shapes or patterns, to monitor the robot's movement, and to supervise the participant's choices and decisions. In these cases, computer vision is referred to as a primary factor in the educational process.

In some other cases, however, computer vision participates as a support activity. The combination of computer vision and educational robotics techniques helps present the course or the traditional activities differently from the teaching chair's stereotype, stimulating the student's interest. In these cases, we consider that computer vision operates as an assistive technology. Table 1 summarizes whether the relevant computer vision activity serves either as a primary factor or as an assistive technology.

Table 1. Classifying computer vision as a primary educational tool or as an assistive technology.

	Primary Educational Tool	Assistive Technology
Altin et al. (2014)	●	
Wu et al. (2019)	●	
Madhyastha (2016)	●	
Kusumota et al. (2018)	●	
Rios et al. (2017)	●	
He et al. (2017)	●	
Amanatiadis et al. (2017)	●	
Amanatiadis et al. (2020)	●	
Jimenez et al. (2019)	●	
Olvera et al. (2019)	●	
Darrah et al. (2018)	●	
Kaburlasos et al. (2018)	●	

Table 1. *Cont.*

	Primary Educational Tool	Assistive Technology
Efthymiou et al. (2020)	●	
Vrochidou et al. (2018)		●
Majgaard et al. (2015)		●
Alemi et al. (2014)		●
Tozadore et al. (2017)	●	●
Karalekas et al. (2020)	●	●
Evripidou et al. (2021)	●	●
Vega et al. (2018)	●	●
Park and Lenskiy (2014)	●	●

As observed by the analysis of articles in this section, computer vision tasks are used as educational tools to help and support both, students and educators. Overall, these 21 papers under the first research question analysis indicated that incorporating computer vision tasks as an education tool in educational robots is valuable for students to build knowledge better and enhance their academic success and/or professional skills. Besides that, the combination of humanoid robots and computer vision helps students increase their interest, enhance their communication skills, and improve their social abilities. As summarized in Table 2 and Figure 5, the topics and areas of interest vary. Computational thinking/programming as well as the playing/interacting with robots are on top positions while STEM and language teaching follow.
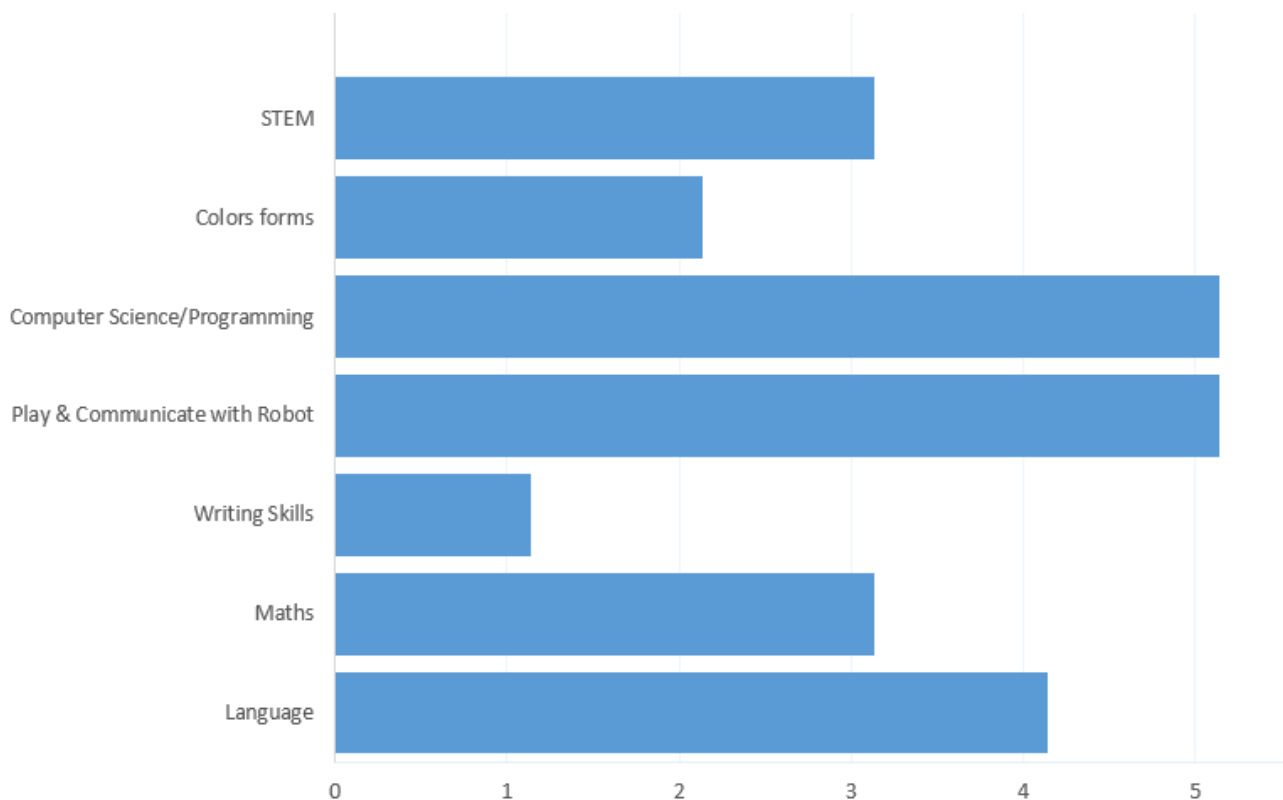

Figure 5. Areas of interest of the articles analyzed.

3.2. How Computer Vision Benefits Educational Robotics' Expected Learning Outcomes in K-12 Education?

The second research question aims to extract the benefits of using computer vision in K-12 education by examining the evidence reported in other research studies. The enhanced capabilities of educational robots, comparing to traditional methods, are the main reason that robotic activities can improve the teaching process. However, the empirical evidence of the impact of robots in education is considered limited in many cases [8]. Giving more intelligence to robots, is one of the future challenges in the design of robotics [51]. The following questions arise: Is computer vision one of the parameters that can help in this direction? Are the expected educational robotics outcomes, as described earlier, contributed through computer vision?

The study conducted in [28] presents how students can boost their learning skills when interacting with a proposed educational robotic system. In the proposed method, the children tried their best to make the robot understand their writing, and then the robot was either correcting or rewarding for their effort. The work conducted in [29] presents a robotics-based system that uses object recognition for teaching English to pre-school children in China. Outcomes showed that a computer vision robot could keep children's interest in learning while improving its efficiency. In both, the previous papers, self-efficacy and motivation skills are enhanced with computer vision.

The authors in [44], present how students are motivated by a robot's presence in the class. The study revealed that the robot's presence increases the engagement during the course, enhances the understanding of mathematical concepts, increases computational and logical thinking, and improves children's cognitive skills. Humanoid robot with computer vision contributes to the Computational thinking and motivation skills in educational robotics.

Moreover, the RALL platform proposed by the authors in [45], attempts to evaluate human robots' use and effectiveness in a game-based learning activity. Experimental outcomes present that RALL students' improvement can triple as they scored higher than non-RALL students. This demonstrates the effectiveness of the RALL system over traditional methods both in short-term and long-term learning. Creativity and problem-solving skills are supported by computer vision in the educational learning process.

According to [36], results suggest that robot-assisted treatment can improve children's behavior. The use of NAO social robots with imitation games can increase children's social and communication skills. Subsequently, the study demonstrated by the authors in [11] investigates how technology can support and enrich the learning process to help students learn more efficiently. Teachers highlighted robots' opportunities that can be used to support children's active experiments. Results show that pupils can quickly become self-propelled, and they can have an excellent academic discussion amongst them. Self-efficacy can be supported with computer vision, as in both articles robots were providing encouraging comments to the children. In addition, in Amanatiadis et al. [38] additional skills such as interaction skills, joint attention, and cognitive flexibility, team spirit, etc. can be enhanced as multi-robot collaborative games can assist children treatment and participation in a social environment with other children. It is worth noting that this is one of the two articles from the examined literature that directly impacts collaboration skills boosted through computer vision tasks.

The authors in [47] present the use of EUROPA robots for STEM teaching and highlight that students can better understand real-world problems. As authors mentioned within the paper, 'The students were acquainted with more advanced technological subjects and were motivated for independent learning and discovery. All students could follow, understand, and work on the EUROPA robots without any serious problems. Some of them were even willing to drill down to the robot's architecture'. Problem-solving and motivation skills are enhanced through computer vision tasks.

Moreover, the strategy proposed by the authors in [40] demonstrated how children could learn the geometrical shapes and colors correctly. Results present that game-like teaching is more attractive to children as they do not get bored. This demonstrates that problem-solving and creativity skills are enhanced with the use of computer vision tasks.

In addition, the authors in [49] outline STEM teaching to children in secondary schools using a PiBot robot as an educational tool supported by a camera. The proposed approach introduces to students computer vision and allows the creation of different exercises that combine vision and vision-based behaviors to practice. In [49], problem-solving skills are supported by computer vision tasks.

Experimental results demonstrated in [39] highlight that an interaction between a human–robot and a user can be achieved quickly. This can aid the children to gain knowledge through activities that use color recognition algorithms. According to the authors . . . 'NAO can hold their attention a little bit longer so the kid can analyze, understand and learn what NAO is saying, increasing the ability to relate the names with the colors, not only those of the figures but the colors that the child sees around him'. Hence, the problem-solving skills are improved with the use of computer vision.

Furthermore, Ref. [41] presents an Open-Source Robotic system with an on-board camera that conducts essential computer vision functions such as shape and color detection. The robot is emphasized in STEM technology and, more specifically, in computer networks. Experimental results show that it can provide students with rich learning experiences. According to the authors, computational thinking and problem-solving skills are boosted through computer vision functions.

The following articles highlight the participant's interaction with humanoid social, creative robots benefiting through play-full environments supported by computer vision. Firstly, the work outline in [32] presents how students reacted when playing with a Cozmo robot. Students quickly understood how to interact with Cozmo, reported their impressions, and indicated that the proposed system could be a useful educational tool. During the drawing process, all students were focused only on Cozmo, so after the robot finished the drawing, the students showed a surprised reaction and applauded it. Secondly, the work conducted in [34] indicates that an educational robot system with contextual teaching characteristics that can mine knowledge from the real world can dramatically improve the enjoyment and engagement of robotic learning. The work presented in [42] outlines the interaction between a NAO robot and children when playing games. Experimental outcomes discussed in the paper are encouraging since the robot learns using crowd-computing feedback techniques. Besides the previous articles, in [46] young age children interacted excitingly with a NAO robot regardless of the method used (autonomous or teleoperated). Authors observe no significant difference between the conditions in the user's enjoyment and time response, and children lessen their perception of the robot's intelligence after learning more about the teleoperation. However, most children (80%) said that they preferred to interact with an autonomous robot. The above articles' main findings indicated that creativity is enhanced using computer vision tasks in play-full games.

Subsequently, the experimental outcomes discussed in [43] present that through the integration of multiple robots, sensors, and modalities, and with the use of an unconstrained and autonomous child–robot system people can achieve a high level of understanding. In the previous work, children felt comfortable playing and communicating with robots, and they believe that robots can behave like humans. The authors reported very characteristically about the multi-party game 'Form a Farm': In case of a wrong guess, the robot reveals more characteristics of the animal (animal color, number of legs, animal class, e.g., mammals, reptiles). In case of correct identification of the animals, the robot asks the children to properly place the animal in a farm with some distinct segmented areas which appears in a touch screen in front of them, aiming to entertain, educate, but also establish a natural interaction between all parties. Experimental results showed that most children (27/31) stated that they like playing with the robots, while 22 enjoyed the play since robots understood both their movements and speech. During the games, computer

vision tasks supported all the tasks mentioned above to create a proper framework for multi-modal communication between children and robots, as it happens between humans. According to the play-full learning environment created by educational, social robots, creativity educational robotics learning outcome is supported by computer vision task. Furthermore, the collaboration learning outcome can be enhanced with computer vision, as mentioned before, since, through interactive games, students communicated and worked together as a team to win the robots.

Furthermore, authors in [50] outline the educational environment's improvement and students' engagement in the learning process by utilizing a hamster robot platform and computer vision algorithms. Using this combination, students were motivated and passionate about programming, and their creativity skills were improved. As mentioned in this paper, creativity, motivation, and computational thinking are enhanced by computer vision tasks as students' firstly extract the robot's position and orientation and then label the obstacles close to the robot. This information is employed to help children make decisions about the robot's future movements.

In [30], authors present a prototype language tutor robot that teaches new words in French and Spanish. Authors reported that the use of various techniques in Machine learning, computer vision, and Speech Processing has helped build a reasonably robust robot tutor which attempts to mimic a human teacher. Experimental results show that participants were overall satisfied by how the robot teaches. The participants evaluated different criteria, including comfort level of communication, the fluidity of the interaction, robustness of individual components, quality of overall experience, and user-friendliness with the robot. The average score from the participants was around 8/10. Furthermore, the participants share their opinions on certain qualities that human teachers have which may be hard to be replaced entirely with a score of 6.7. The MonitoRE system reported by the authors in [33] presents that students feel motivated and demonstrate more interest in educational robotics which considers the practicality in computer vision tasks. More precisely, authors reported . . . 'students felt more motivated, demonstrating interest in using monitored task environments because it eases the understanding of the difficulties the moving robot faces in completing the activities, assisting students in the teaching-learning process'. In a sample of 46 users, 93% said that they found it interesting and enjoyed the experience, 86% noted that they were satisfied with the usability of the proposed system and the established scenarios, 95% considered the correction of the proposed efficient design, and 84% reported that they had obtained a learning return with the proposed approach. By analyzing both previous papers, we concluded that motivation skills are improved with computer vision.

The results obtained in [48] present that participants can increase their algorithmic/programming thinking skills while developing a positive attitude towards programming. Outcome results show that the hybrid group has rated its experience satisfaction high and requires less average time to complete the exercises than students who attended the entire course and other students who completed all activities. Motivation and self-efficacy skills are boosted with computer vision as children received encouraging messages from the robot. Moreover, computational thinking and problem-solving skills are also enhanced through the robot's advice to children.

Overall, under the spectrum of the second research question, the literature highlights that the implementation of computer vision techniques and tasks in educational robotics appears to benefit the overall teaching/educational process. Based on the observations analyzed in detail earlier, computer vision tasks increase students' interest in learning and motivate them to search for something new. Moreover, the formation of appropriate computer vision tasks into traditional education robotics activities improves the participants' social and communication skills and helps them better understand and then solve real-world problems.

More precisely, computer vision efficacy on the six expected learning outcomes of educational robotics are highlighted for each relevant paper and are summarized in Table 2. Furthermore, the hypothesis about humanoid and Social robots relationship with computer vision in educational robotics extracted from the 'computer vision in educational robotics' index related terms from the bibliographic map (Figure 3) is confirmed since 60% of robotics platform argue this assessment as can be seen in Table 3.

Table 2. Correlating the expected learning outcomes of educational robotics with computer vision.

	Problem Solving Skills	Self Efficacy	Computat. Thinking	Creativity	Motivation	Collaborat.
Altin et al. (2014)		●			●	
Wu et al. (2019)		●			●	
Madhyastha (2016)					●	
Kusumota et al. (2018)				●		
Rios et al. (2017)					●	
He et al. (2017)				●		
Amanatiadis et al. (2017)		●				
Amanatiadis et al. (2020)		●				●
Jimenez et al. (2019)	●					
Olvera et al. (2018)	●			●		
Darrah et al. (2018)	●		●			
Kaburlasos et al. (2018)				●		
Efthymiou et al. (2020)				●		●
Vrochidou et al. (2018)			●			
Majgaard et al. (2015)		●				
Alemi et al. (2014)	●			●	●	
Tozadore et al. (2017)				●		
Karalekas et al. (2020)	●				●	
Evripidou et al. (2021)	●	●	●		●	
Vega et al. (2018)	●					
Park and Lenskiy (2014)			●	●	●	

Table 3. Summarizing the context of related literature.

Paper	Age/Level	Area Explored/Topic(s)	Computer Vision	Robot Used
Altin et al. (2014)	Young children	Writing Skills	Image Processing	NAO
Wu et al. (2019)	Pre-school children	English as a second language	Object Recognition	Kinect, Projector, PC
Madhyastha (2016)	From 14+ ages	French or Spanish	Image Processing Object Recognition	A prototype robot language tutor
Kusumota et al. (2018)	Elementary Children	Mathematical operations, spelling, directions, and questions functions	Image and Textual Processing	Cozmo
Rios et al. (2017)	Secondary School	Teaching Robotics	Color and Image processing	MonitoRE
He et al. (2017)	Pre-school children	Metacognition tutoring and geometrical thinking training the pronunciation of the vocabulary in both Chinese and English	Image processing Object Recognition	NAO
Amanatiadis et al. (2017)	Children diagnosed with Autism Spectrum Condition (ASC)	Social communication and interaction skills, joint attention, response inhibition	Color Recognition and Image Processing	NAO
Amanatiadis et al. (2020)	ASC	Social communication and interaction skills, joint attention, response inhibition	Color Recognition and Image Processing	2 NAOs
Jimenez et al. (2019)	Children diagnosed with Syndrome Down	Colors Forms	Color Recognition	NAO
Olvera et al. (2018)	Pre-school, 3 and 4 years	Geometric Forms and Colors	Object and Color Recognition	NAO
Darrah et al. (2018)	Secondary Children	learning of computational thinking and STEM, with an emphasis on Computer Science concepts	Shape and Color detection	Open-Source Robotics
Kaburlasos et al. (2018)	From 5 to 14 years old	Play and Communicate with Robots	Image Processing and Pattern Recognition	NAO
Efthymiou et al. (2020)	Elementary, 6 to 11 years	Play and Communicate with Robots	Image Processing Object recognition, Distance Speech Recognition	ChildBot (NAO, Furhat, Zeno)
Vrochidou et al. (2018)	K12 system (ages 16–17)	Mathematics	Camera Presenting	NAO

Table 3. *Cont.*

Paper	Age/Level	Area Explored/Topic(s)	Computer Vision	Robot Used
Majgaard et al. (2015)	Between 11 and 16 years old	Programming, language learning, ethics, technology and mathematics	Camera Presenting	NAO
Alemi et al. (2014)	Children at the age of 12	English Vocabulary	Camera Presenting	NAO
Tozadore et al. (2017)	Elementary Between 7 and 11 years	Interact with robot	Camera Presenting/ object recognition	NAO
Karalekas et al. (2020)	K12 system (ages 16–17)	STEM Teaching in sciences, engineering and programming	Using camera for watching and as a color sensor	EUROPA
Evripidou et al. (2021)	Pre-school children	Programming/algorithmic thinking	Camera Presenting/ Fuzzy system	Bee-bot
Vega et al. (2018)	Pre-university children in secondary schools	STEM Teaching	Color Recognition, Visualisation through camera	PiBot
Park and Lenskiy (2014)	Secondary School	Language Programming	Camera with computer vision algorithms	Hamster robot

3.3. How Affordable and Feasible Is the Integration of Solutions That Combine Educational Robotics and Computer Vision in K-12 Instructional Activities?

The third research question aims to present if the integration of computer vision in education can be applied from a cost–benefit perspective and robot models' availability.

Educational robotic platforms are available in a wide range and vary in cost, parts, and complexity [51]. To provide thorough answers to the third research question, we investigated the following aspects: (i) the design specifications and how complex it is to build a robot? Furthermore, (ii) what is the cost of the various types of robots' outlined in the literature?

Initially, based on the data summarized in Table 3, 60% of relevant papers uses NAO in educational activities. NAO is an up-and-coming robotic system with tremendous potential that incorporates computer vision to affect the learning process, as presented during the analysis of the second research question. However, it is still not affordable to educators as its cost is too high [51].

According to [47], the Europa robot model significantly cheaper compared to NAO. The overall cost of EUROPA robot is estimated less that 120 euros. EUROPA is 'a two-wheel, inexpensive differential drive robot with a manipulator' easy to be integrated into the educational learning process. EUROPA is adequately scalable and flexible to fit into different educational levels and curricula, and it allows introductory or advanced-level programming, depending on educational level. In addition, the authors in [41] present an attractive alternative solution to the most expensive kits commercially available that can aid the instruction of multiple STEM + computer science related topics. This study's proposed approach aims to increase computational thinking strategies among high school computer science course students.

The work conducted in [50] outlines a low-cost implementation based on a sensor simulation to test their proposed computer vision algorithm. Authors claim that the selected hamster robot development platform is inexpensive compared to others (around 150 euro). Moreover, [30] describes a low-cost robot language tutor with interaction capabilities that can be personalized. At a broad system level, the main components used in this prototype robot include a microphone to capture speech, a camera that moves in different directions and captures frames, and of course, a processing unit.

The proposed system developed and presented in [29] which is still in an early stage of development, explores new teaching methods for pre-school children. It consists of three components that can be easily procured, a projector, a Kinect, and a computer system, making it possible to be applied in a school. The task environment recommended by the authors in [33] is composed of artificial landmarks, including the mobile robot, and allows the monitoring system to identify and evaluate the overlap of colors and shapes established in the environment. In this way, teachers and students can use any educational robotic kit during the learning process, depending on the cost efficacy.

Furthermore, the study conducted in [32] presents the development and implementation of an educational platform for the Cozmo mobile robot. As the authors mentioned, the proposed educational tool is a low-cost solution (around 150 euro) that can be applied in the education environment, providing teachers the ability to create and run their scripts. In [49], the authors claim that the PiBot platform is a low cost (under 180 euros) solution for supporting STEM teaching which makes it affordable for most schools and students. Apart from the price, the robot hardware can be easily assembled using a 3D printer that follows the Do It Yourself (D.I.Y.) philosophy. As mentioned by the authors, the 3D PiBot model and all the developed plugins are. The Bee-pot is an easy to use, low-cost educational platforms (around 75 euro) that can be integrated into pre-schools to support the learning procedure. Childbot as mentioned by [43], combines three different robotic platforms (NAO, Furhat and Zeno), available in the market but their integration into the educational process climbs to their high cost (about 39,000 euro).

To sum up, as analyzed in this subsection, the final research question instigated that most of the robotic platforms presented in this study are applicable to be integrated into the

educational process. In this study, it can be observed that Robotic models are available in a wide range, consist of a low-cost solution, and can be quickly developed in K-12 education. More precisely, 60% of Robotics model in relevant articles adopt the humanoid social robot NAO as presented in Table 3. NAO is a very promising robot capability, according to the literature, but its high cost (about 13,000 euro), may be seen as an obstacle to its integration into the educational process.

4. Conclusions

To highlight the synergies and the intersections between educational robotics and computer vision and demonstrate how the combination of these disciplines impact K-12 education, this paper maps all relevant research studies using a systematic mapping process.

This study aims to present how robotics autonomy gained through computer vision supports educational robotics by examining three significant factors. Firstly, to identify the role of computer vision in educational robotics. Secondly, to determine the benefits of computer vision in educational robotics, and thirdly investigate how efficient it is to apply computer vision in K-12 education. After a systematic search in online bibliographic databases using keyword searching and a snowballing approach, we extracted and analyzed 21 primary articles from the recent literature.

Based on the performed analysis, computer vision-related tasks in educational robotics demonstrate high potential in teaching assistance. Children's gain in learning is significant as determined by the selected articles' outcomes analysis. In the comparison groups, it was found that in those who were assisted with the computer vision procedures, the participants demonstrated more interest in the educational process, learned the concepts they taught more easily, spent less time completing their work, and generally were very satisfied with the way of teaching. The results highlighted that the most common use of computer vision in educational robotics is as a primary factor for teaching while, a limited number of studies (only 3) presented computer vision as an assistive tool only. Regarding the discussion in all relevant articles, the results were positive about computer vision tasks' effectiveness to support the learning process. The research claimed the increase in academic achievement from pre-school to secondary schools and special education in different subjects area and skills, as summarized in Figure 5.

Moreover, through relevant articles, computer vision correlation with the six expected learning outcomes of educational robotics was presented, and the results are summed up in Table 2. It is noteworthy that 'Creativity', 'Motivation', and 'Problem-Solving Skills' are considered the most common learning outcomes supported by computer vision activities involving educational robotics, with 'Self-Efficacy' and 'Computational Thinking' to follow. The 'Collaboration' learning outcome appears with two degrees of participation; the articles' analysis has shown that the interactive game's use in the learning process involves all students to cooperate and develop a team spirit to succeed in a mission or to win. Future research directions in this area are needed to create an educational process that supports this outcome. One example of this direction is computer vision analysis of students' behavior and interaction during group tasks on how they communicate, teach others, and how comfortable they are with fellow students. Later, it enhances peer-to-peer interaction between students as per their comfort levels.

Besides, computer vision integration in schools depends on the robot model's availability and the cost factor. From the relevant article's analysis, we found that all the robot models presented in this study can easily enhance the educational process, and they are available for K-12 education. However, 60% of the documents used the humanoid and social NAO robot, recognizing its many potentials, without considering the high integration cost.

To sum up, computer vision-related tasks in educational robotics are considered useful tools for the learning process. The convergence of computer vision and educational robotics is still in the incipient phase. It is worth exploring ways in which such technology can

mutually benefit the students involved in the education process to develop the proposed outcomes skills by making learning more effective, garner more attention from them, maximize their interest, customizing courses and materials as per their understanding capabilities, and most importantly, fun.

According to the different perspective observed by the systematic analysis, the essential factors that influence educational robotics enhanced with computer vision tasks in K-12 education effectiveness, include usability and availability of appropriate learning activities and content (knowledge area to be explored), children age group, robot models to be used and cost parameter. Possible applications to be designed must consider the robot as communication mediators to support group learning, interacting with the robot in a playful environment, the children can respond with high motivational levels and creativity, focus on children interests and weakness could improve self-efficacy, and helps to problem-solving and computational thinking skills. Further researchers would help develop more applications (design new or modify current robotic activities) to use computer vision in educational robotics.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Goswami, S.; Uddin, M.S.; Islam, M.R. Implementation of Active Learning for ICT Education in Schools. *Int. J. Innov. Sci. Res. Technol.* **2020**, *5*, 455–459. [CrossRef]
2. Hegedus, S.; Moreno-Armella, L. Information and communication technology (ICT) affordances in mathematics education. *Encycl. Math. Educ.* **2020**, 380–384. [CrossRef]
3. Qi, Y.; Pan, Z.; Zhang, S.; van den Hengel, A.; Wu, Q. Object-and-Action Aware Model for Visual Language Navigation. *arXiv* **2020**, arXiv:2007.14626.
4. Hong, Y.; Wu, Q.; Qi, Y.; Rodriguez-Opazo, C.; Gould, S. A Recurrent Vision-and-Language BERT for Navigation. *arXiv* **2020**, arXiv:2011.13922.
5. Scaramuzza, D.; Achtelik, M.; Doitsidis, L.; Fraundorfer, F.; Kosmatopoulos, E.B.; Martinelli, A.; Achtelik, M.W.; Chli, M.; Chatzichristofis, S.A.; Kneip, L.; et al. Vision-Controlled Micro Flying Robots: From System Design to Autonomous Navigation and Mapping in GPS-Denied Environments. *IEEE Robot. Autom. Mag.* **2014**, *21*, 26–40. [CrossRef]
6. Kouskouridas, R.; Amanatiadis, A.; Chatzichristofis, S.A.; Gasteratos, A. What, Where and How? Introducing pose manifolds for industrial object manipulation. *Expert Syst. Appl.* **2015**, *42*, 8123–8133. [CrossRef]
7. Evripidou, S.; Georgiou, K.; Doitsidis, L.; Amanatiadis, A.A.; Zinonos, Z.; Chatzichristofis, S.A. Educational Robotics: Platforms, Competitions and Expected Learning Outcomes. *IEEE Access* **2020**, *8*, 219534–219562. [CrossRef]
8. Benitti, F.B.V. Exploring the educational potential of robotics in schools: A systematic review. *Comput. Educ.* **2012**, *58*, 978–988. [CrossRef]
9. Toh, L.P.E.; Causo, A.; Tzuo, P.W.; Chen, I.M.; Yeo, S.H. A review on the use of robots in education and young children. *J. Educ. Technol. Soc.* **2016**, *19*, 148–163.
10. Piaget, J. Part I: Cognitive development in children: Piaget development and learning. In *J. Res. Sci. Teach.* **1964**, *2*, 176–186. [CrossRef]
11. Majgaard, G. Multimodal robots as educational tools in primary and lower secondary education. In Proceedings of the International Conferences Interfaces and Human Computer Interaction, Las Palmas de Gran Canaria, Spain, 22–24 July 2015; pp. 27–34.
12. Shoikova, E.; Nikolov, R.; Kovatcheva, E. Conceptualizing of Smart Education. *Electrotech. Electron. E+ E* **2017**, *52*.
13. Savov, T.; Terzieva, V.; Todorova, K. Computer Vision and Internet of Things: Attention System in Educational Context. In Proceedings of the 19th International Conference on Computer Systems and Technologies, Ruse, Bulgaria, 13–14 September 2018; pp. 171–177.

14. Zhong, B.; Xia, L. A systematic review on exploring the potential of educational robotics in mathematics education. *Int. J. Sci. Math. Educ.* **2020**, *18*, 79–101. [CrossRef]

15. Angel-Fernandez, J.M.; Vincze, M. Towards a definition of educational robotics. In Proceedings of the Austrian Robotics Workshop, Innsbruck, Austria, 17–18 May 2018; p. 37.

16. Gabriele, L.; Tavernise, A.; Bertacchini, F. Active learning in a robotics laboratory with university students. In *Increasing Student Engagement and Retention Using Immersive Interfaces: Virtual Worlds, Gaming, and Simulation*; Emerald Group Publishing Limited: West Yorkshire, UK, 2012.

17. Misirli, A.; Komis, V. Robotics and programming concepts in Early Childhood Education: A conceptual framework for designing educational scenarios. In *Research on e-Learning and ICT in Education*; Springer: New York, NY, USA, 2014; pp. 99–118.

18. Alimisis, D. Educational robotics: Open questions and new challenges. *Themes Sci. Technol. Educ.* **2013**, *6*, 63–71.

19. Mubin, O.; Stevens, C.; Shahid, S.; Mahmud, A.; Jian-Jie, D. A Review of the Applicability of Robots in Education. Available online: http://roila.org/wp-content/uploads/2013/07/209-0015.pdf (accessed on 1 January 2013).

20. Jin, L.; Tan, F.; Jiang, S. Generative Adversarial Network Technologies and Applications in Computer Vision. *Comput. Intell. Neurosci.* **2020**, *2020*, 1459107. [CrossRef]

21. Klette, R. *Concise Computer Vision*; Springer: Cham, Switzerland, 2014.

22. Szeliski, R. *Computer Vision: Algorithms and Applications*; Springer Science & Business Media: Cham, Switzerland, 2010.

23. Boronski, T.; Hassan, N. *Sociology of Education*; SAGE Publications Limited: London, UK, 2020.

24. Bebis, G.; Egbert, D.; Shah, M. Review of computer vision education. *IEEE Trans. Educ.* **2003**, *46*, 2–21. [CrossRef]

25. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Group, P.; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* **2009**, *6*, e1000097. [CrossRef]

26. Petersen, K.; Feldt, R.; Mujtaba, S.; Mattsson, M. Systematic mapping studies in software engineering. In Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering (EASE), Bari, Italy, 26–27 June 2008; pp. 1–10.

27. Yli-Huumo, J.; Ko, D.; Choi, S.; Park, S.; Smolander, K. Where is current research on blockchain technology?—A systematic review. *PLoS ONE* **2016**, *11*, e0163477. [CrossRef]

28. Altin, H.; Aabloo, A.; Anbarjafari, G. New era for educational robotics: Replacing teachers with a robotic system to teach alphabet writing. In Proceedings of the 4th International Workshop Teaching Robotics, Teaching with Robotics & 5th International Conference Robotics in Education, Padova, Italy, 18 July 2014; pp. 164–166.

29. Wu, Q.; Wang, S.; Cao, J.; He, B.; Yu, C.; Zheng, J. Object recognition-based second language learning educational robot system for chinese preschool children. *IEEE Access* **2019**, *7*, 7301–7312. [CrossRef]

30. Madhyastha, M.; Jayagopi, D.B. A low cost personalised robot language tutor with perceptual and interaction capabilities. In Proceedings of the 2016 IEEE Annual India Conference (INDICON), Bangalore, India, 16–18 December 2016; pp. 1–5.

31. Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; Darrell, T. Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM international conference on Multimedia, Mountain View, CA, USA, 18–19 June 2014; pp. 675–678.

32. Kusumota, V.; Aroca, R.; Martins, F. An Open Source Framework for Educational Applications Using Cozmo Mobile Robot. In Proceedings of the 2018 Latin American Robotic Symposium, 2018 Brazilian Symposium on Robotics (SBR) and 2018 Workshop on Robotics in Education (WRE), João Pessoa, Brazil, 6–10 November 2018; pp. 569–576.

33. Rios, M.L.; Netto, J.F.d.M.; Almeida, T.O. Computational vision applied to the monitoring of mobile robots in educational robotic scenarios. In Proceedings of the 2017 IEEE Frontiers in Education Conference (FIE), Indianapolis, IN, USA, 18–21 October 2017; pp. 1–7.

34. He, B.; Xia, M.; Yu, X.; Jian, P.; Meng, H.; Chen, Z. An educational robot system of visual question answering for preschoolers. In Proceedings of the 2017 2nd International Conference on Robotics and Automation Engineering (ICRAE), Shanghai, China, 29–31 December 2017; pp. 441–445.

35. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv* **2015**, arXiv:1506.01497.

36. Amanatiadis, A.; Kaburlasos, V.G.; Dardani, C.; Chatzichristofis, S.A. Interactive social robots in special education. In Proceedings of the 2017 IEEE 7th International Conference on Consumer Electronics-Berlin (ICCE-Berlin), Berlin, Germany, 3–6 September 2017; pp. 126–129.

37. Chatzichristofis, S.A.; Boutalis, Y.S. CEDD: Color and Edge Directivity Descriptor: A Compact Descriptor for Image Indexing and Retrieval. In *Computer Vision Systems, Proceedings of the 6th International Conference, ICVS 2008, Santorini, Greece, 12–15 May 2008*; Lecture Notes in Computer Science; Gasteratos, A., Vincze, M., Tsotsos, J.K., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5008, pp. 312–322. [CrossRef]

38. Amanatiadis, A.; Kaburlasos, V.G.; Dardani, C.; Chatzichristofis, S.A.; Mitropoulos, A. Social robots in special education: Creating dynamic interactions for optimal experience. *IEEE Consum. Electron. Mag.* **2020**, *9*, 39–45. [CrossRef]

39. Jiménez, M.; Ochoa, A.; Escobedo, D.; Estrada, R.; Martinez, E.; Maciel, R.; Larios, V. Recognition of Colors through Use of a Humanoid Nao Robot in Therapies for Children with Down Syndrome in a Smart City. *Res. Comput. Sci.* **2019**, *148*, 239–252. [CrossRef]

40. Olvera, D.; Escalona, U.; Sossa, H. Teaching Basic Concepts: Geometric Forms and Colors on a NAO Robot Platform. *Res. Comput. Sci.* **2019**, *148*, 323–333. [CrossRef]

41. Darrah, T.; Hutchins, N.; Biswas, G. Design and development of a low-cost open-source robotics education platform. In Proceedings of the 50th International Symposium on Robotics (ISR 2018), Munich, Germany, 20–21 June 2018; pp. 1–4.

42. Kaburlasos, V.; Bazinas, C.; Siavalas, G.; Papakostas, G. Linguistic social robot control by crowd-computing feedback. In Proceedings of the JSME annual Conference on Robotics and Mechatronics (Robomec), Kyushu, Japan, 2–5 June 2018; p. 1A1-B13.

43. Efthymiou, N.; Filntisis, P.P.; Koutras, P.; Tsiami, A.; Hadfield, J.; Potamianos, G.; Maragos, P. ChildBot: Multi-Robot Perception and Interaction with Children. *arXiv* **2020**, arXiv:2008.12818.

44. Vrochidou, E.; Najoua, A.; Lytridis, C.; Salonidis, M.; Ferelis, V.; Papakostas, G.A. Social robot NAO as a self-regulating didactic mediator: A case study of teaching/learning numeracy. In Proceedings of the 2018 26th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 13–15 September 2018; pp. 1–5.

45. Alemi, M.; Meghdari, A.; Ghazisaedy, M. Employing humanoid robots for teaching English language in Iranian junior high-schools. *Int. J. Humanoid Robot.* **2014**, *11*, 1450022. [CrossRef]

46. Tozadore, D.; Pinto, A.; Romero, R.; Trovato, G. Wizard of oz vs. autonomous: Children's perception changes according to robot's operation condition. In Proceedings of the 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Lisbon, Portugal, 28 August–1 September 2017; pp. 664–669.

47. Karalekas, G.; Vologiannidis, S.; Kalomiros, J. EUROPA: A Case Study for Teaching Sensors, Data Acquisition and Robotics via a ROS-Based Educational Robot. *Sensors* **2020**, *20*, 2469. [CrossRef] [PubMed]

48. Salomi, E.; Amanatiadis, A.; Christodoulou, K.; Chatzichristofis, S.A. Introducing Algorithmic Thinking and Sequencing using Tangible Robots. *IEEE Trans. Learn. Technol.* **2021**. [CrossRef]

49. Vega, J.; Cañas, J.M. PiBot: An open low-cost robotic platform with camera for STEM education. *Electronics* **2018**, *7*, 430. [CrossRef]

50. Park, J.S.; Lenskiy, A. Mobile robot platform for improving experience of learning programming languages. *J. Autom. Control Eng.* **2014**, *2*, 265–269. [CrossRef]

51. Pachidis, T.; Vrochidou, E.; Kaburlasos, V.; Kostova, S.; Bonković, M.; Papić, V. Social robotics in education: State-of-the-art and directions. In Proceedings of the International Conference on Robotics in Alpe-Adria Danube Region, Patras, Greece, 6–8 June 2018; Springer: Cham, Switzerland, 2018; pp. 689–700.