

The role of computer vision in the development of knowledge-based systems for teaching and learning of English language education

George Chibuikwe Agbo^{1*} and Philomina Akudo Agbo²

Lecturer, Department of Computer & Robotics Education, University of Nigeria, Nsukka¹

Lecturer, Department of Arts Education, University of Nigeria, Nsukka²

Received: 04-May-2020; Revised: 28-May-2020; Accepted: 29-May-2020

©2020 George Chibuikwe Agbo and Philomina Akudo Agbo. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

This study focused on the assessment of the use of computer vision in the teaching and learning of English language education. The use of computer vision in the teaching and learning of language education has become a novel approach that has attracted varying attentions in recent times. In this study, descriptive analysis was used to provide a coherent view of the use of computer vision in implementing knowledge-based systems for the teaching and learning of English Language Education. The language education tasks that were analysed comprised of image captioning, human-machine interaction, video captioning, visual attributes, visual retrieval and visual question answering. The study also looked at the approaches that could be used to incorporate computer vision and the teaching and learning of English language education courses, with particular reference to models as an integrated topic of distributional semantics. The researchers made an analysis of computer vision and English Language teaching and learning, using distributional semantics as both words and imagery. An integrated view of the subject matter was clearly presented and useful suggestions were made for promising future guidelines. The study also x-rayed the usefulness of computer vision for effective teaching and learning of language education and suggested novel approaches that could be used in incorporating computer vision in developing knowledge-based system for effective teaching and learning of the English language.

Keywords

Computer vision, English language education, Teaching and learning, Images.

1. Introduction

The teaching and learning of the English language are perhaps one of the most complex fields of human learning; however, computer vision has become an indispensable tool, both within and outside the classroom. According [1], computer vision is the use of images and videos to describe ideas and real-life situation. Also stated that computer vision is the presentation and interpretation of ideas through the use of images and videos [2]. Noted that the importance of computer vision in teaching and learning cannot be over emphasized [3]. Also noted that although the computer vision is not here to replace human beings, its impact on teaching and learning are enormous [4]. The author further argued that computer vision would be able to extend human capabilities and possibilities of teaching, learning, and research in the field of language learning when effectively utilized.

In the same vein, [5] noted that the field of computer vision has recently grown such that it can help teachers to teach more effectively.

The author further stressed that the use of computer vision in language education would aid teachers in differentiating instructions, so that learners can receive more tailored teaching. The key motivation for using computer vision in language teaching and learning is to support the development of systems that could help teachers, as well as learners in maximizing learning [6]. Noted that this involves understanding and modeling learners and teachers, as well as effective pedagogies and learners' contexts, which can be more effective through the use of computer visions for the development of knowledge-based systems [7].

Knowledge-based systems have become crucial visual aids in teaching, learning and research. A knowledge-based system is referred to as computer software that is designed to behave akin to a human

*Author for correspondence

expert. Defined knowledge-based system as computer software that imitates the decision-making ability of a human expert [8]. In the same vein, [9] argued that a knowledge-based system is a developed computer program that emulates the way that human experts perform different tasks in specific fields. Similarly, [10] noted that knowledge-based systems are an intelligent computer program that uses the knowledge gotten from human beings to solve setting problems. Also noted that knowledge-based systems are systems that store human knowledge, so that users can interact with them in order to solve problems in specific fields, such as the teaching and learning of language education [11].

Language Education can be viewed as the processes involved in the teaching and learning of languages, such as English language. Defined language education as the process imbibing in the learner the ability to communicate with a specific language [12]. Stated that, English language education refers to the ways through which English language is taught in schools [13]. Opined that English language education is a field of learning that has to do with the teaching and learning of English language [14]. Pointed out that English language education deals solely with equipping learners with the degree of knowledge they need so as to effectively use a language. Therefore, English language education is a type of teaching and learning that helps the students to develop in them the competences required for the use of English languages [15].

In English language education, the role of computer vision in facilitating teaching and learning of language education cannot be overemphasized. According to [16], the roles of computer vision can be summarized with the use of three key terms, namely; reorganization, recognition, and reconstruction (3Rs), which is otherwise referred to as the concept of the 3Rs. The author further noted that reconstruction involves the use a various processes to bring information together from different viewpoints, while recognition involves the use of labels to connect different object and parts in a vision or image, while reorganization involves the representation of the structure of an idea in a sequential manner, ranging from low-level vision to high-level tasks, such as semantic segmentation. Opined that computer vision helps to clearly understand the grammatical structure of a sentence, as well as provide a more informative description of images in a heap of unorganized words [17]. According to [18], computer vision helps to paint

some imagery that often comes with accompanying text which can be used in the teaching and learning of different aspects of the English language, such as verb agreement.

In developing a knowledge-based system for the teaching and learning of the English language, certain parameters can be adopted. Stated that in developing knowledge-based systems, computer visions could be used to generate images that would be able to describe grammar and sentence information needs to be extracted [19]. Considering parts of speech in English language teaching and learning, the quadruplets of Nouns, as well as verbs and prepositions can be used to represent meaning that is created from visual detectors in a knowledge-based system [20]. Further noted that visual modules create various items which may be an object or subject in a given sentence [21]. Opined that numerous models such as, Hidden Markov Model can be used to decipher the major likely sentences from a limited combination of four related ideas, along with some strategies that describes images by predicting the most likely verb and scene (preposition) during the development of knowledge-based systems for the teaching and learning of language education [22]. went further to explain that computer vision can be used in capturing and accessing language learning in a non-obstructive manner without continuous efforts from humans, such as the use of computer vision to learn the use of adjectives in sentences using a language model and generated as a template-based sentence in a knowledge-based system [23, 24]. Thus, the main objective of the study is to ascertain the sole of computer vision in the development of knowledge-based systems for teaching and learning of English language education. Specifically, the study seeks to:

1. Identify how the 3Rs (reconstruction, recognition and reorganization) can be designed in computer vision to connect ideas for the teaching and learning of the English language.
2. Understand how the ideas that are presented in computer vision can be translated into human readable texts for the teaching and learning of the English language?

Research Questions

The two research questions that were used for the study are:

1. How can the 3Rs be designed in computer vision to connect ideas for the teaching and learning of the English language?

2. How can the ideas that are presented in computer vision be translated into human readable texts for the teaching and learning of the English language?

2. Methodology

The study adopted exploratory research design. According to [25] exploratory research design is a type of research design that usually depends on direct observation of a small selection of what is to be studied. Exploratory research design is therefore

considered most appropriate for this study because the researcher explored the use of images in establishing traditional computer vision tasks in a system and gives the output/result inform of a meaningful learning object, which could be used in the development of knowledge-based systems for teaching and learning of English language education. A diagram of the flow of the method can be seen in *Figure 1*.

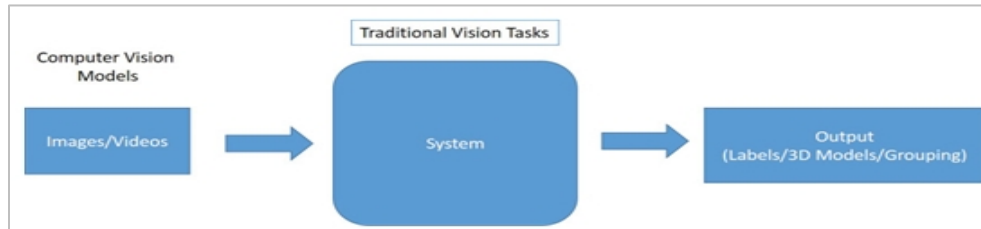


Figure 1 Diagram of the flow of this research work

3. Results

The design of the 3Rs in computer vision for connecting ideas for the teaching and learning of English language

The tasks of computer vision were summarized through the use reconstruction, recognition, and reorganization, which is otherwise referred to as the concept of the 3Rs. In the computer vision constructed, a man was seen cutting a red apple. Numerous interpretations could be given to an image by organizations, which could be recognized from a cartoon viewpoint of object models that are in a restructured manner. This means that the representations of semantic can equip learners with useful information from English language to assist in providing a clear understanding of the events, objects, relations and actions in different teaching and learning situations.

The reconstruction in this study was accomplished using various processes through the incorporation of useful information from different and relevant viewpoints, with the help of the three-dimensional (3D) objects that led to the production of a single and meaningful visual image. The reconstruction was carried out using depth images, which resulted in a 3D model. The recognition and reorganization of the 3D model were done through the assignment of labels to objects in the images. In support of this process, [26] opined that 3Rs has an overlapping contribution to knowledge-based design using computer vision, because it helps in the recognition of related tasks. According to [27] a useful scenario can be segmented using both low-level and high-level vision information, such as shadow segmentation that utilizes class information which can be used for effecting teaching and learning. The design of the 3Rs in computer vision for connecting ideas for teaching and learning can be seen in the *Figure 2*.

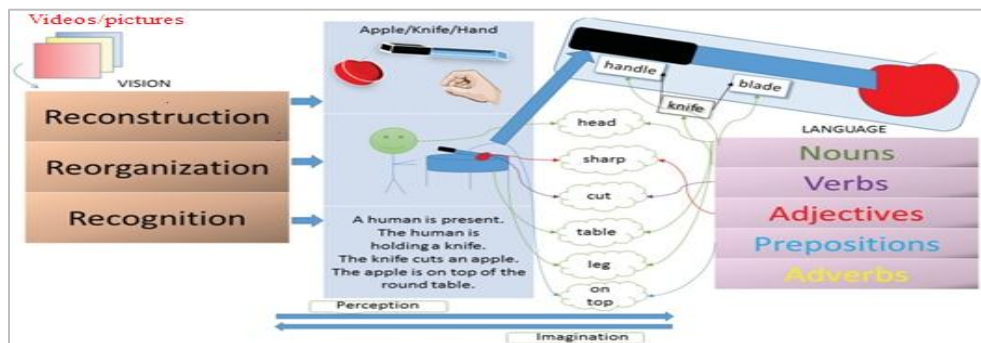


Figure 2 The design of 3Rs in computer vision for connecting ideas for teaching and learning

How the ideas that are presented in computer vision be used as a tool for making predictions on the follow-up task in the lesson, such as logical inference, semantics recognitions and compositional semantics grammar. In this regard, the syntax may include, but not limited to the composition of smaller units of English language, such as words to larger units, such as sentences and phrases (compositionality) and the study of word forms (morphology). In support of this, [29] stated that computer vision can be used in building up a link from visual data to language data, such as words or phrases, in the form of labels that are attached to image patches that contains an object with a word as object recognition. The author further noted that labelling a background in an image is scene recognition and assigning words for pixel grouping is semantic segmentation. The ideas that are presented in computer vision and translated into human readable texts can be seen in *Figure 3*.

In the development of a knowledge-based system using computer vision, machine translation triangle proposed by [28] can be adopted. In the case of teaching and learning of English language education topics such as syntax, semantics and discourse, different activities can be summarized into concepts which can be grouped from syntax to semantics, discourse and to pragmatics at the top level, so as to achieve effective communication. From the figure 2 below, it can be evident that, this design can concentrate on the important players, as well as the objects and activities which could be considered of particular event as its interest, choices, attention, as well as awareness mechanisms, which may involve, English language education and vision, respectively. The system could be

used as a tool for making predictions on the follow-up task in the lesson, such as logical inference, semantics recognitions and compositional semantics grammar. In this regard, the syntax may include, but not limited to the composition of smaller units of English language, such as words to larger units, such as sentences and phrases (compositionality) and the study of word forms (morphology). In support of this, [29] stated that computer vision can be used in building up a link from visual data to language data, such as words or phrases, in the form of labels that are attached to image patches that contains an object with a word as object recognition. The author further noted that labelling a background in an image is scene recognition and assigning words for pixel grouping is semantic segmentation. The ideas that are presented in computer vision and translated into human readable texts can be seen in *Figure 3*.

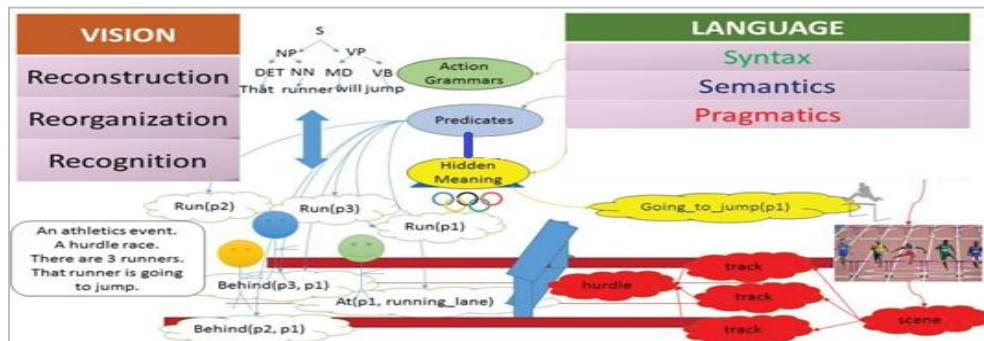


Figure 3 Computer vision and translated into human readable texts

4. Limitations of the study

The following are the limitations of the study:

1. Although there this study has summarized the tasks of computer vision using the 3Rs, there are still other numerous dimensions of decisions making using computer vision which.
2. The images/computer vision produced in this work is based on a single topic in language education. This constitutes a limiting factor to the generalization of the findings in other topics/areas of English language learning.

5. Suggestions for further research

The following related areas have been suggested for further research

1. The present study can be replicated using field of learning such as mathematics.
2. A similar study that will cover motion analysis and noise reduction techniques in computer vision.

3. Another study on computer vision and English language learning can be carried out using survey research design.
4. A study can also be conducted on identifying the constraints militating against the successful use of computer vision in the teaching and learning of computer vision.

6. Conclusion

The continued emergent of technology has brought a paradigm shift in the field of teaching and learning from the traditional methods to information and technology driven forms which necessitates for the use of computer visions in developing knowledge-based systems for teaching and learning of specific subject matters or courses such as English language education. The question then is: what does it actually mean for a knowledge-based system to understand and interpret images for English language education? And how does a human being understand what is presented in computer vision? These questions gave rise to the idea

of assessing the role of computer Vision in the development of knowledge-based systems for Teaching and Learning of English Language Education. This research has succeeded in suggesting useful frameworks which could be useful in understanding the ideas behind the use of images in English language teaching and learning. This was x-rayed in both the *Figures 1* and *2* above.

The figures also showed that, in developing a knowledge-based computer vision, the system should have the capability of recognizing and interpreting labels, different tools and actions, such as the direction of arrows, interaction between objects, as well as human and their parts, depending on the subject matter being dealt with. In translating to textual form for better comprehension, the system is designed in such a way that it should have the capability of answering questions which are related to the reason behind what, why and how the vision relates and interacts to each other. This also means that the system should be capable of recognizing the intentions of the individual key players, as well as in predicting the possibility of an event occurring in the sequence. More so, in the application or use of the computer vision for teaching and learning, questions should be asked about the characteristics of an event and how several events are interrelated to one another, while the synthesis should ascertain various ways that can be used to construct new events, drawing from the ideas of existing phenomena. In carrying out an evaluation, there is need to ascertain the level of utility of each event.

Acknowledgment

None.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

- [1] National Science Foundation (US). Directorate for Education, Human Resources. Shaping the future: New expectations for undergraduate education in science, mathematics, engineering, and technology. National Science Foundation, Division of Undergraduate Education; 1996.
- [2] Coppola D. Integrating teaching and research. *ASEE Prism*. 1997; 7(4):18-22.
- [3] National Research Council (US). National committee on science education standards, assessment. National science education standards: draft for review and comment only. National Research Council; 1994.
- [4] Shah M, Bowyer K. Mentoring undergraduates in computer vision research. *IEEE Transactions on Education*. 2001; 44(3):252-7.

- [5] Shah M, Bowyer K. Mentoring undergraduates in computer vision research. *IEEE Transactions on Education*. 2001; 44(3):252-7.
- [6] Maxwell BA. A survey of computer vision education and text resources. *International Journal of Pattern Recognition and Artificial Intelligence*. 2001; 15(05):757-73.
- [7] Sanchez A, Velez JF, Moreno AB, Esteban JL. Introducing algorithm design techniques in undergraduate digital image processing courses. *International Journal of Pattern Recognition and Artificial Intelligence*. 2001; 15(05):789-803.
- [8] Grewe L. Effective computer vision instruction through experimental learning experiences. *International Journal of Pattern Recognition And Artificial Intelligence*. 2001; 15(05):805-21.
- [9] Pridmore TP, Hales WM. Understanding images: an approach to the university teaching of computer vision. *Engineering Science & Education Journal*. 1995; 4(4):161-6.
- [10] Glasgow N. A guide to student-centered, problem-based learning. 1997.
- [11] Duch BJ. Problem-based learning in undergraduate education. 2019.
- [12] Novins K, McCane B. Incorporating primary source material into the undergraduate computer vision curriculum. *International Journal of Pattern Recognition and Artificial Intelligence*. 2001; 15(05):775-87.
- [13] Donelan M, Wallace J. Peer assisted learning: a truly co-operative initiative. *Students Supporting Students*. 1998:11-22.
- [14] Ploetzner R, Dillenbourg P, Preier M, Traum D. Learning by explaining to oneself and to others. *Collaborative learning: Cognitive and Computational Approaches*. 1999; 1:103-21.
- [15] Jiménez-Peris R, Khuri S, Patiño-Martínez M. Adding breadth to CS1 and CS2 courses through visual and interactive programming projects. In the proceedings of the thirtieth SIGCSE technical symposium on computer science education 1999 (pp. 252-6).
- [16] Fell HJ, Proulx VK. Exploring Martian planetary images: C++ exercises for CS1. In proceedings of the twenty-eighth SIGCSE technical symposium on computer science education 1997 (pp. 30-4).
- [17] Astrachan O, Rodger SH. Animation, visualization, and interaction in CS 1 assignments. *ACM SIGCSE Bulletin*. 1998; 30(1):317-21.
- [18] Montgomery RE. Image analysis: a group assignment in programming with breadth. In proceedings frontiers in education 1995 25th annual conference. engineering education for the 21st Century 1995 Nov 1 (pp. 4d3-12). IEEE.
- [19] Stockman G, Enbody R. Teaching advanced students C++ with computer vision. In workshop on combined research-curriculum development in computer vision, Kauai, Hawaii 2001.
- [20] Marks J, Freeman W, Leitner H. Teaching applied computing without programming: a case-based introductory course for general education. In

proceedings of the thirty-second SIGCSE technical symposium on computer science education 2001 (pp. 80-4).

- [21] Egbert D, Bebis G, McIntosh M, LaTourette N, Mitra A. Computer vision research as a teaching tool in CS1. In 32nd annual frontiers in education 2002 (pp. T4G-T4G). IEEE.
- [22] Egbert D, Bebis G, Williams D. Computer vision research teaching modules for community college computer science and engineering courses. Age. 2003; 8:1.
- [23] Sarkar S, Goldgof D. Integrating image computation in undergraduate level data-structure education. International Journal of Pattern Recognition and Artificial Intelligence. 1998; 12(08):1071-80.
- [24] <http://www.inf.unibz.it/~nutt/Teaching/DSA1415/dsa-assignments.html>. Accessed 26 march 2020.
- [25] Shirish TS. Research methodology in education. Lulu.com; 2012.
- [26] Hoover A. Computer vision in undergraduate education: Modern embedded computing. IEEE Transactions on Education. 2003; 46(2):235-40.
- [27] Fink E, Heath M. Image-processing projects for an algorithms course. International journal of pattern recognition and artificial intelligence. 2001; 15(05):859-68.
- [28] Alhawiti KM. Natural language processing and its use in education. Computer Science Department, Faculty of Computers and Information technology, Tabuk University, Tabuk, Saudi Arabia. 2014.
- [29] Wiriyathamabhum P, Summers-Stay D, Fermüller C, Aloimonos Y. Computer vision and natural language processing: recent approaches in multimedia and robotics. ACM Computing Surveys (CSUR). 2016; 49(4):1-44.



Dr. George Chibuiké Agbo is Lecturer, Department of Computer & Robotics Education, University of Nigeria, Nsukka. He holds a Doctor of Philosophy (P.hD) in Computer & Robotics Education, A Master of Technology (M.Tech) in Information Systems Management, a Bachelors of Science (B.Sc) and Diploma in Computer Education. He possesses over 30 International IT Certifications from Europe and has published over 15 local and international journal articles.

Email: chibuiké.agbo@unn.edu.ng



Philomina Akudo Agbo, Lecturer at Department of Arts Education, University of Nigeria, Nsukka. She holds a Master of Education (M.Ed) in Language Education (English), a Bachelors of Arts Education (BA.Ed) in English Language Education. She received a merit award as the best graduating student of the department of Arts Education, University of Nigeria, Nsukka and currently a Ph.D student of the same department. She has published over 5 local and international journal articles.