Aisha Rahmayanti, 2023, 11:4
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Utilizing Deep Learning Algorithm for Enhanced Visual Product Identification

Aisha Rahmayanti, Saliyah Kahar,Trismayanti Dwi P
[1]Management and Science University,Malaysia
[2]Management and Science University,Malaysia
[3]State Polytechnic of Jember, Indonesia
aisharahmayanti10@gmail.com,saliyah_kahar@msu.edu.my,trismayanti@polije.ac.id

Abstract-Deep learning algorithms have revolutionized the field of computer vision in recent years. One such area is visually impaired individuals, who rely on visual product identification apps to navigate the world around them. One particularly effective, cost-efficient supervised learning approach that may be used to solve a variety of challenging problems is deep learning. Deep learning makes it possible to achieve better results because it has a variety of illustrative qualities and does not rely on any specific learning techniques. Deep learning is widely employed in many applications, including image classification, face identification, speech recognition, visual recognition, language processing, object detection, and numerous science and business analysis. deep learning is explained through a thorough examination of deep learning architectures, their traits, and their drawbacks. Many studies have developed the assistive technology using android application which useful in today world and people with visually impaired can make use of this technology to help them accomplished on their daily routine. CNN and other deep learning methods can be utilized in the development of visual product identification algorithms. This training process involves repeated iterations of providing input data and adjusting algorithm parameters to improve product identification capabilities. For image recognition tasks, CNN are frequently utilized, including visual product identification. They are designed to automatically learn and extract features from images by using multiple layers of convolution and pooling operations. CNN have achieved modern performance in a variety of image recognition tasks.

Keywords- Deep Learning Algorithm; Visually Impaired Identifying Products

I. INTRODUCTION

The people that are visually impaired or blind do not really use smartphones because it is fully touchscreens and not many ways for them to command their phones for help other than the basic, which is calls as that is possible through the help of Google Assistant or SIRI for Apple or iOS users. Thus, it is hard to go shopping with smartphone applications for the blind or visually impaired persons. Furthermore, persons who are blind or visually impaired cannot use almost all the applications that are in a smartphone as there is no way to use it without touching the screen or bare minimal touches and use of voice command. Currently, In the market there are many apps, but no application that can help a visually challenged person by telling him or her what is available and what is not available in a specific store as soon as they scan a certain take a picture of what do they want. People who are blind or visually handicapped deal with numerous issues every day. These people struggle to distinguish and identify the objects around them, therefore they rely solely on direction,

International Journal of Science, Engineering and Technology

An Open Access Journal

especially for daily tasks. The most challenging for visually impaired are that the ability of them to recognize the color, shape and differentiate the currency of the money.Nowadays, there are many of assistive technology that can help them as a sighted guidance and improves the quality life of visually impaired (Smys et al., 2020)(Vargas, 2017)

The visually impaired mostly now have the smartphone as it become a basic necessity of each individual. Thus, it provides a great platform to develop an application specifically for visually impaired to assist them. Found that the person with visual impairments are frequently used the mobile application for their daily activities (Nasir et al., 2021). Furthermore, they are looking for some improvement and new application that can help them to less dependent on others. Specific mobile task as such landmark detection using the mobile phone. His findings become the platform on designing the technology that facilitates visually impaired (Ganesan et al., 2022). Due to its vast advantages in picture classification, deep learning is becoming more significant.(Voulodimos et al., 2018), applications for language and speech processing, data mining (Roslan et al., 2019). The system now beats earlier learning models thanks to the incredible rise of data and technological advancements, which have brought about a number of critical pathways in deep learning models.

Models for object prediction (Bakator & Radosav, 2018), Prior to applying a classification approach, earlier products were built on the basis on extracting hand-engineered qualities.

Later, the process was streamlined by the reformation of neural networks (RNN) in 2012 and the introduction of cutting-edge architectures including the pyramid system, Retina-Net networks, You Only Look Once (YOLO), Single Shot Multi-Box Detector (SDD), and Convolution Neural Network (CNN). Despite its high efficiency, wearable device implementation is challenging due to the maximum processing expenses. As a result, individuals who are blind or visually handicapped use portable gadgets to forecast objects. As a result, the researchers have developed an economical and effective method that can forecast various items. However, it is still difficult to determine an object's precise location (Ooi & Ibrahim, 2021).

II.OBJECTIVES OF THE STUDY

The objective of this study is threefold. Firstly, it aims to evaluate the advancements in deep learning algorithms used in visual product identification applications and assess their recognition capabilities. This evaluation will help determine the extent of improvement achieved through these new algorithms. Secondly, the research seeks to propose an effective deep learning algorithm specifically designed to enhance visual product identification for individuals with visual impairments. The goal is to provide a solution that enables visually impaired individuals to easily identify and recognize products using technology. Finally, the study aims to validate the proposed deep learning algorithm by developing a prototype. The prototype will serve as a practical implementation of the algorithm, allowing for real-world testing and evaluation of its effectiveness in improving visual product identification.

1. Significance of the Study

The accessibility of visual product identification apps for visually impaired individuals has brought about a remarkable transformation in their lives. These apps empower them to independently navigate and interact with the products in their environment. With the help of these innovative tools, visually impaired individuals can now identify various products and gain a sense of independence and autonomy. They can make informed decisions while shopping, select desired products, and efficiently manage their daily activities. This newfound ability not only enhances their overall quality of life but also fosters a greater sense of self-reliance. The development of these apps, particularly those employing deep learning algorithms, signifies a significant technological breakthrough. It highlights the potential of such algorithms to address real-world challenges and demonstrates how artificial intelligence can profoundly improve accessibility for individuals with visual impairments. Through this advancement, the possibilities of AI are showcased, opening doors for further innovations in the field of accessibility technology.

Deep Learning Algorithm for Visual Product Identification

Deep learning algorithms have shown great success in visual product identification tasks. Commonly used deep learning algorithms for visual product identification, For image identification applications,

International Journal of Science, Engineering and Technology

An Open Access Journal

CNN are frequently utilized, including visual product identification. They are designed to automatically study and extract features from images by using multiple layers of convolution and pooling operations. CNN has achieved advanced performance in many image recognition tasks. A common deep learning algorithm used for identifying visual products other than CNN.

III.SCOPE AND LIMITATIONS

The proposed new deep learning algorithm will be developed specifically for visual product identification in mobile applications for the visually impaired. The algorithm will utilize image recognition and classification techniques to accurately identify and provide audio feedback on products through a smartphone camera. The algorithm will be developed and tested on a limited range of products commonly found in retail settings. The accuracy of the algorithm may be affected by poor lighting or low-quality camera hardware on certain mobile devices. Additionally, the algorithm may not be effective in identifying products that have similar visual features or packaging.

IV.MATERIALS AND METHODS

Deep learning algorithms have shown great success in visual product identification tasks. Commonly used deep learning algorithms for visual product identification, For image identification applications, CNN are frequently utilized, including visual product identification. They are designed to automatically study and extract features from images by using multiple layers of convolution and pooling operations. CNN has achieved advanced performance in many image recognition tasks. A common deep learning algorithm used for identifying visual products. CNN automatically learn hierarchical features from raw images through the use of convolutional filters applied across the image. These filters capture local patterns, edges, and textures at different spatial scales.

Scale Invariant Feature Transform (SIFT)

SIFT is an algorithm for extraction and description of features in digital images. The SIFT algorithm works by identifying important features in an image that are reliable and invariant to scale, rotation and shift. These features can be used to compare and match different images, as well as to perform object recognition in images.

CNN, or Convolutional Neural Networks :

CNN for problems involving image recognition, including visual product identification. They are designed to automatically learn and extract features from images by using multiple layers of convolution and pooling operations. CNN have achieved modern performance in a variety of image recognition tasks.

RNN : Recurrennt nerural networks

RNN are often used for order modeling assignments, such as processing of natural language and speech. They can also be applied to visual product identification tasks by treating the image as a sequence of pixels. One popular RNN based approach is the a network called the Long Short-Term Memory (LSTM) that may simulate temporal dependencies in the image features.

V.RESULTS AND DISCUSSION

CNN use convolutional layers to learn filters that extract features from input images, followed by pooling layers to down sample the feature maps. CNN have been widely applied in image recognition tasks, such as identifying objects in images, facial recognition, and medical image analysis.Then RNN have a feedback loop that enables them to pass information from previous time steps to the current time step. This allows RNN to capture temporal dependencies in the input data. RNN have been applied in numerous tasks involving natural language processing, including text generation, speech synthesis, and sentiment analysis. The last SIFT detects distinctive points in an image, called key points, and describes them using a set of descriptors that are invariant to changes in scale, rotation, and illumination. SIFT has been widely applied in image matching, object recognition, and 3D reconstruction.

Table 1. Comparison of Deep Learning Algorithm

Differentiation Component	Deep Learning Algorithm		
	SIFT	CNN	RNN

Architecture	It identifies and describes local characteristics in an image using a traditional computer vision technique. using a scale-space approach. It involves finding key point locations, estimating their orientations, and extracting feature descriptors.	It is a type of deep learning algorithm specifically designed for image analysis tasks. CNN use multiple convolution layers to automatically learn spatial hierarchies of features from input images.	This type of deep learning algorithm is used for sequential data processing, such as natural language processing or time series analysis. RNN have recurrent connections that allow them to retain memory of past inputs.

Feature Extraction	SIFT algorithm explicitly detects and extracts local features (key points) that are invariant to changes in scale, rotation, and illumination. It generates a feature descriptor for each key point.	CNN automatically learn hierarchical features from raw images through the use of convolution filters applied across the image. These filters capture local patterns, edges, and textures at different spatial scales.	RNN operate on sequential data and capture temporal dependencies. They can learn to extract features from sequential inputs by processing them one element at a time, while retaining a memory state.
Training and Learning	SIFT is a handcrafted feature extraction method and does not involve explicit training. The algorithm computes scale-space extreme and extracts local features based on a set of predefined rules.	CNN are trained in a supervised manner, where the model learns to classify or regress based on labeled training examples. The training process involves forward and backward propagation through the network's	RNN can be trained using various techniques such as back propagation through time (BPTT) or variants like long short-term memory (LSTM) and gated recurrent units (GRU).

International Journal of Science, Engineering and Technology

An Open Access Journal

		layers.	
Applications	SIFT is commonly utilized for computer vision tasks like object recognition, image matching, and 3D reconstruction. It is robust to changes in scale, rotation, and affine transformations.	CNN excel in tasks involving images, such as picture segmentation, object detection, and classification, and image generation. They have been instrumental in achieving state-of-the-art performance in these areas.	RNN are widely employed in tasks involving natural language processing, including sentiment analysis, machine translation, language modeling, and speech recognition. They can effectively handle sequential and variable-length inputs.

SIFT is a classical computer vision approach, CNN are well-suited for image analysis tasks, and RNN are commonly used for sequential data processing and natural language processing tasks. Each of these algorithms has its own strengths and applications in different domains of computer vision and machine learning.

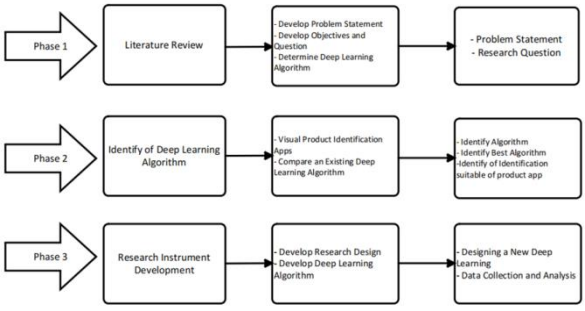

Fig. 1. Research Methodology

The research process comprises three phases. In Phase one, an extensive literature review was conducted by exploring various journals focusing on deep learning algorithms. This allowed for an in-depth understanding of the subject, identification of problem statements, and formulation of clear objectives and research questions. Moving into Phase two, the research involved a comparative analysis of different deep learning algorithms to determine the most suitable one for product identification. Several algorithms were assessed to find the best fit for the specific requirements. Finally, Phase three involved implementing a selected deep learning algorithm. This phase encompassed instrument development, research design, and the creation of a new deep learning model. Additionally, data collection was carried out, followed by rigorous analysis to evaluate the algorithm's performance and its suitability for the intended purpose.

VI.CONCLUSION AND RECOMMENDATION

In conclusion, the app has the potential to greatly benefit visually impaired individuals by providing them with improved access to information and increased independence. Its innovative features, such as text recognition and object identification, have shown promising results during testing and user feedback. However, there are areas where further improvements and updates can be made to enhance the overall user experience and address certain challenges.

REFERENCES

[1] Bakator, M., & Radosav, D. (2018). Deep learning and medical diagnosis: A review of literature. Multimodal Technologies and Interaction, 2(3). https://doi.org/10.3390/mti2030047

[2] Ganesan, J., Azar, A. T., Alsenan, S., Kamal, N. A., Qureshi, B., & Hassanien, A. E. (2022). Deep Learning Reader for Visually Impaired. Electronics (Switzerland), 11(20), 1–22. https://doi.org/10.3390/electronics11203335

[3]Nasir, H. M., Brahin, N. M. A., Aminuddin, M. M. M., Mispan, M. S., & Zulkifli, M. F. (2021). Android based application for visually impaired using deep learning approach. IAES International Journal of Artificial Intelligence, 10(4), 879–888. https://doi.org/10.11591/ijai.v10.i4.pp879-888

[4] Ooi, Y. K., & Ibrahim, H. (2021). Deep learning algorithms for single image super-resolution: A systematic review. Electronics (Switzerland), 10(7). https://doi.org/10.3390/electronics10070867

[5] Roslan, M. F. M., Shaiful, A. I. M., Beleed, A. E.,

International Journal of Science, Engineering and Technology

An Open Access Journal

Omar, M. N. B., & Illias, S. (2019). Paper of the Title 2. 2(X), 1–5. https://doi.org/10.35940/ijrte.xxxxx.xxxxxx

[6] Smys, S., Iong, J., & Chen, Z. (2020). Survey on Neural Network Architectures with Deep Learning. 02(03), 186–194.

[7] Vargas. (2017). Deep Learning : a Review Deep Learning : a Review. Advances in Intelligent Systems and Computing, July. https://doi.org/10.20944/preprints201810.0218.v1

[8] Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. Computational Intelligence and Neuroscience, 2018. https://doi.org/10.1155/2018/7068349

Author's details

Aisha Rahmayanti, Management and Science University, Malaysia, aisharahmayanti10@gmail.com
Saliyah Kahar, Management and Science University, Malaysia, saliyah_kahar@msu.edu.my
3Trismayanti Dwi Puspitasari, State Polytechnic of Jember, Indonesia, trismayanti@polije.ac.id