

A Comparative Review on Object Detection System for Visually Impaired

Dr K Sreenivasulu^a, P. Kiran Rao^b, Dr.VenkataRamana Motupalli^c

^aProfessor of CSE,G.Pullaiah College of Engineering and Technology ,Kurnool,Email:- sreenu.kutala@gmail.com

^bAssociate Professor in CSE,G.Pullaiah College of Engineering and Technology Kurnool kiranraocse@gmail.com

^cAssociate professor in CSE,Annamacharya Institute of Technology and Sciences,Kadapa ,Email:- venkataramana_558@yahoo.co.in

Article History: Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

Abstract: Vision is one of the key senses allowing citizens to communicate with the natural world. There are about two hundred million blind people globally and visually disabled people obstruct numerous everyday practices. It is also really critical that blind people recognize their world and realize with which items they communicate. This paper review all the method and tool related to camera-based device to enable the blind person interpret text patterns written on items kept in hand. This is the system for helping individuals with visual disability interpret and translate text patterns to the audio output. The framework first suggests the approach to take an image from the camera and the area of the target to retrieve the object from the context and derive a text pattern from that object. Different algorithm is assessed in various scenes. The observed text is linked to the blueprint and translated into the performance of the voice. Localized and binarized text patterns utilising Optical Character Recognition (OCR). The text is translated to an audio output. The voice quality is given to the blind person.

Keyword: OCR, Visually blind person, Machine learning, Deep Learning, SVM, Accuracy

1. Introduction

Computer science has taken a fundamental role in the development of the daily activities of the human being, presenting tools that provide solutions to problems in different areas. There are lots of research focuses on artificial intelligence; emphasizes in his machine learning course how applications based on bio-inspired algorithms, machine learning and evolutionary techniques, allow for example to have information about traffic, make weather predictions, generate security with biometric recognition, control crops, obtain location thanks to automatic mapping or even allow us to interact on social networks.

The challenge of computer science is to extract useful information from the environment in which humans interact, in order to create mathematical, statistical or quantitative models that can represent these natural processes of man [1]. From there we try to put all these techniques at the service of people to facilitate fluid interaction with texts, images and conversations. It is also at this point where it has been shown that there are certain barriers that prevent said fluid interaction, for example there are physical limitations that make people have problems both in receiving information and in communicating ideas (blind or deaf-mute people), on the other hand there may be cultural barriers,

Therefore, the motivation of this research is given, the fact of building a technological tool supported by computer science (in this case Deep Learning) that allows to overcome some of these barriers, through the creation of a service that recognizes and automatically characterizes images taken or provided by a user.

The following text presents in the background, a historical tour of the evolution of computer science and, more precisely, the evolution that probabilistic algorithms have had, as well as works that likewise have wanted to identify objects in images using deep learning.

A. Background

Taking into account that this is a work focused on developing a technological tool based on computer science, it is really important to contextualize its birth, some of its history and development in the service of human beings, and then deepen the techniques that they are planned to be implemented in the development. Computer science has been historically recorded since the construction of the first useful devices to keep accounts and solve mathematical problems. Over the years there were important contributions from researchers such as Leibniz, Pascal and Babbage to achieve an approximation to a first computer and the first algorithms.

The development of computers was then limited to the advancement of new technologies, the machines that were created were the size of rooms. Still, downsizing these machines wasn't the only important issue; scientists were looking for a way to make these machines increasingly intelligent [2]. That is why the investigations of relevant people in history began to appear, characters like Alan Turing considered the father of computer science, who began to make an abstraction of the human brain to represent it in the world of computers, understanding that in this way the machines could be made better not only in the hardware part, but also in the software part.

It was then Messrs. Walter Pitts and Warren McCulloch, a neurophysiologist and a mathematician respectively, who conceived the foundations of neural computing, modeled in 1943 a simple neural network using electrical circuits [3]. From this moment on, it became practically a new line of research to represent the functioning of the brain in mathematical techniques and models applied to computer computers.

In 1980, the decade considered to be the age of illustration in computer science, Ray Solomonoff, inventor of algorithmic probability, built the foundations for the beginning of a really important and fundamental computational technique in computer applications currently used, machine learning "Machine Learning", an area whose purpose is to create programs capable of generalizing behaviors from unstructured information supplied in the form of examples. This technique was complemented over the years with topics that were developed in parallel, including all the knowledge acquired from neural networks to optimize computational processes. In summary, these machines that are provided as already mentioned, a series of examples with their respective outputs, Companies such as Facebook, Google or YouTube today use learning algorithms to make the interaction of their platforms and their users increasingly intelligent and personalized, taking as reference tastes, customs and recurring activities [4].

In 2006 Geoffrey Hinton, a specialist in cognitive psychology and neural networks establishes the concept of deep learning "Deep Learning", which can be considered an evolution of machine learning, with a similar idea and through more robust algorithms it seeks to process language natural human beings, texts, audios, videos and images for problem solving [5], making the machines look more and more like a person. In the same way as "machine learning", with the aim of creating more complete applications, it was important to link with more computer science techniques [6,7], including artificial vision, data mining, bio-inspired systems, robotics and artificial intelligence.

Image Recognition Based on Deep Learning [8]. This work was carried out by researchers from the Chinese University of Computer Science and Technology in 2015, they apply algorithms of the two most important currents of Deep Learning, convolutional artificial neural networks and deep belief networks, in classification problems of images, with this they manage to determine the high efficiency of these two models and of course they can compare their performance, this work leads them to the conclusion that convolutional neural networks have a superior performance than their competition. **Hybrid Deep Learning for Face Verification** [9]. In this research developed in 2016, it was sought to make a hybrid between a convolutional neural network algorithm, and the Boltzman machine algorithm. The main objective was to carry out a study for the verification of faces, starting from the extraction of local visual characteristics of the face compared in two images, these data are processed through multiple layers.

Pose and Category Recognition of Highly Deformable Objects Using Deep Learning [10]. In this work, the researchers focus their efforts on a highly complex problem such as the recognition of easily deformable objects, in this case garments hung from a single point, the system was designed and implemented in a robotic platform in charge of manipulating the garment and data extraction from it, achieving a high degree of performance in the classification task. This research shows the versatility and high efficiency of convolutional networks applied to digital images.

Convolutional-Recursive Deep Learning for 3D Object Classification [11]. These researchers from Stanford University (including Andrew Ng). They took advantage of the advances and new techniques of artificial vision to determine depth in images, with the aim of classifying objects in RGB images with a model based on the combination of convolutional neural networks and recursive neural networks. To carry out the tests, a database with 51 kinds of household objects was obtained. With 300 examples for each class, where each one of them was seen from three different angles to make the training with a more robust base.

The previous articles show how the solution to problems of detection and classification of objects in videos or images, are oriented towards the use of convolutional neural networks thanks to the high performance rates shown and the use or splicing that occurs from probabilistic techniques and artificial vision techniques.

A. Justification

Technology seeks to transform the environment to meet the needs of man, sometimes the needs arise due to the existence of barriers that limit their capabilities, these barriers can be cultural, regional, intellectual or even physical. This degree project seeks to eliminate some of the limitations created by these barriers by developing a tool that, by applying computational intelligence techniques, is capable of giving those who use it a better understanding of their environment and in certain cases improve their quality of life.

On the other hand, autonomous learning techniques are one of the most relevant topics today. Hundreds of researchers work daily for the improvement of these techniques, which is why it has been decided to apply this recent and growing technique to implement a useful tool for humans.

Learning and machine vision techniques have not reached their maximum development, there are still no algorithms that are capable of solving any type of problem no matter what it is, in addition, machine learning techniques are quite new, the first machine learning projects They date from the eighties, and those of more complex techniques based on machine learning such as Deep learning are even more recent, therefore there is an opportunity to contribute in the development of these.

II LITERATURE REVIEW

Deep identification of the neural network (DCNN)[12] of objects is a strong visual perception technique that demands enormous device and communication costs. We have proposed a fast and low-power object recognition

processor that lets visually disabled people understand their environment. We have developed an integrated DCNN quantitative algorithm with 8-bit fixed points that effectively measure data 32 values and uses 5-bit indexes to display them, minimising hardware costs relative to the 16-bit DCNN to a marginal precision. The architecture of a particular hardware accelerator uses configurable process engines to conduct multi-layer pipelines which minimise or significantly remove the temporary off-chip transfer data. A search table is used to apply all multiplications in convolutions to minimise power significantly. The design is constructed of SMIC 55-nm and only 68-mw with a 1,1-v voltage of 155GB/s with a maximum output of 2.2 Top/w after a structure simulation has been done.

This article illustrates the usefulness of an electromagnetic sensor[13] to autonomously move visually impaired and blind individuals. It is known that people with vision disorder appear to be helped by the most common white cane. Our design is a traditional white cane microwave radar that allows customers to be aware of an obstacle in a broader, more safe range.

The technology suggested is better than the existing electronic means of transport, noise tolerance and lower measurements. The following are recent advances with an emphasis on miniaturization of circuit boards and antennas in this research activity. A laboratory prototype has been developed and conducted and the first barrier detection test results prove the effectiveness of the unit.

A number of obstruction identification (OD) methods [14] for the monocular vision were created to help visually disabled individuals. Using the vector differences or the object size of two consecutive frames, most traditional OD methods detect obstacles. However, short-term OD efficiency is greatly impacted by the tracking mistake, which results in unreliable OD results. This paper proposes a new OD approach focused on a new framework named Deformable Grid to tackle this issue (DG). The DG is initially a standard grid format, but may be increasingly deformed, based on the object's movement in the scene. The suggested approach detects the collision danger entity depending on the deformation degree in the DG. Experimental findings reveal that the OD system proposed beats the traditional approach in terms of processing time and precision.

Echolocation allows people with visual impairments or no vision[15] to sense spatial information with reflected echoes. However, this procedure often requires thorough precise planning and echolocation under various conditions. Moreover, those who perform this sensing technique must generate the sound and interpret the audio data collected simultaneously. This paper presents and tests the LIDAR (Lass) spatial sensing and range-assistant system to overcome these limitations by collecting spatial information from the user using a LIDAR sensor and translating spatial information to the stereo sound of various pits. The comparatively stereo sound reflects information about the vertical and horizontal distances of the objects, increasing the spatial perception of the surroundings and potential challenges of visually impaired communities. Step I is the hardware and software setup on the LASS computer and Phase II is an overview of system stability. The Penn State Administrative Review Board approved 18 volunteers from the Department of Psychology in Penn. This paper shows that individuals with blindfolded LASS systems can quantitatively classify exterior barriers, discern their relative distances and differentiate the angular orientation of different objects at low levels.

Because of the exponential growth of mobile technologies, a

technology [16] is evolving which can identify banknotes and coins to help visually disabled citizens utilising smartphone embedded cameras. In previous research, robust features have often been handmade, such as scale-invariant characteristics turn or accelerate, and cannot yield robust identification results for banknotes or corners recorded in complex environments and contexts. With recent developments in deep learning technologies, some banknote and coin recognition experiments have been performed utilising a profound convolutionary neural network (CNN). However, these experiments demonstrated degraded efficiency according to context and climate shifts.

This paper offers a 3-stage identification technology with quick regional CNN, geometric constraints and residual networking for new banknotes and coins (ResNet). The experiment carried out by Jordanian dinar and 6,400 images of 8 types of Korean winning banknotes and coins on our smartphones produced more result than state of the art methods based on manufactured features and profound functionality. The experiment was carried out with a Jordanian Dinar (JOD).

Navigation support for visually disabled persons (NAVI) applies [17] to devices that may aid or direct vision deficiency individuals with sound signals, varying from partly sighted to fully blind. In this article, a new NAVI method focused on visual and range knowledge is introduced. We choose a single unit, a consumer RGB-D camera, instead of using several sensors, and gain from both spectrum and visual knowledge. The key contribution is, in particular, the combination of the depth details and the picture strength, contributing to a rigorous extension of the segmentation of the surface. On the one side, the accurate yet restricted profundity information is improved by the long-range visual information. On the other side, it encourages and enhances the difficult and vulnerable to error picture processing with depth knowledge. The framework proposed detects and classifies the key structural components of the scene, which enable the user to move in an unknown manner. The device has been evaluated in a broad spectrum of conditions and data sets to prove performance and that the framework is stable and operates in demanding indoor conditions.

The creation of electronic sensing devices for visually disabled persons [18] involves awareness of their needs and ability. This paper provides a rough study that can be used to describe the parameters for the construction of these devices properly. The emphasis would be on clear-cut metrics, stressing their position in orientation and mobility activities. A new computer is presented that belongs to this class. The detector is designed on a multisensor technique and uses intelligent signal processing to warn the consumer of the presence of artefacts hindering its trajectory. Experimental experiments show the device's effectiveness.

The paper suggests a modern electronic accessibility cane (EMC) to help visually disabled persons [19] sense barriers and discover methods of doing so. The key feature of this cane is that it produces the logical chart of the system in which priority knowledge can be accessed. It provides a simpler representation of the world without overloading details. This priority information is conveyed to the subject through into- it iveness, audio or voice input. The other developments of the EMC include the identification of staircases and the non- formal distance scaling method. It also gives details on the condition of the floor. It consists of an integrated device of low power with ultrasonic sensors and protection indications. In order to validate its design and test its potential to support participants in their everyday mobility, the EMC was subjected to many clinical assessments. Medical tests were done for 16 participants fully blind and four with poor vision. Both participants went with EMC and conventional white cane in a regulated and real-world test setting. The findings of the assessment and large amounts of subjective tests indicate the utility of the EMC in vision recovery programmes.

A number of smartphone or wearable guidance devices [20] have built in the past decades to support visually disabled individuals in established or unfamiliar indoor or outdoor conditions. Three key types are required for these systems: electronic travel aids (ETAs), electronic guide aids (EOAs) and locator devices (PLDs). This paper provides a comparative survey of the portable/wearable obstacle detection/avoidance devices (a sub-category of ETAs) to educate the testing group and the consumers regarding their strengths and the advances achieved in assistive technology. The survey is focused on different characteristics and output parameters of structures, which categorize them, with qualitative and quantitative measurements.

It is understood that being visually disabled [21] is one of the most daunting experiences in existence, and many individuals are faced with this circumstance. As the fields of computer vision and machine learning have evolved, it has become easier for researchers and engineers to develop a support framework. Two basic elements are software and hardware modules of obstacle detection systems. In recent years, deep learning techniques, open source databases, and programming platforms have been turned into operating systems. There are numerous profound learning and open source libraries and in this project we concentrated on programming the framework with Python and making a classification on a small machine named Raspberry Pi using Tensor flow models. We used the ssdlite mobilnet v2 coco model to detect 9 distinct items that may be located on sidewalks. Since the auditory sense is the key stimulus for visually disabled persons, and should not be silenced, we recommended to alert the individual to a mixture of auditory and tactile senses. A device named eSpeak was used to warn the consumer via head phones to the detected object's name. Simultaneously, three separate vibration sensors were positioned in three different locations to the right, centre and left. The vibration sensor with the object name has been triggered when an obstacle is sensed within one of the predefined bounding boxes.

The goal of this article is to develop an obstacle detection [22] device that supports people with visual disability. A basic flight time camera records real-time photographs found in the subject's environment. This photo is evaluated for depth and location of artefacts using OpenCV and MATLAB. The map of inequalities generated is categorised into deep categories. 3 categories would decide which obstacles are relevant, namely close, medium and far. A matrix of this data is generated that correlates to the environment. A temporary Three Dimensional surrounding model is built with actuators and blocks according to this matrix. The subject then will feel this model on an elevated surface to work the way through the obstacles.

In today's sophisticated high-tech environment [23], the need for independent living is recognised for individuals with visually disabled disabilities who face key social limitation issues. They struggle without some manual help in a strange place. Visual information is the base of certain activities, but individuals with visual disability are at a disadvantage when necessary information regarding the world is not accessible.

With the latest developments in inclusive technologies, assistance for individuals with visual disability will be expanded. This project is intended to support people with artificial intellect, computer learning, picture and text recognition who are blind or visually disabled. This concept is applied through Android's mobile app focusing on voice assistant, picture recognition, currency recognition, e-books, talk bot, ... The software will allow you to identify items in the immediate environment by utilising a voice order, conduct text analysis to recognise the text in the hard copy paper. It would also be an effective way for blind people to communicate with the world with computers and to allow use of technology services.

This paper reveals a camera-based device that allows the blind to read text patterns written on handheld items [24]. This is a device that allows people with visual difficulty to understand and convert text patterns into the audio performance. The first step in the framework is to take a picture from the camera and the goal region to recover the object from its context and to extract a text pattern from that object. The safest text is observed with

full healthy external regions (MSER). A new algorithm is tested in multiple scenes. The text observed is related to the strategy and converted into voice output. Located and binarized text patterns for the identification of optical character (OCR). The text is turned into an audio output. The standard of the voice is provided to the blind person. MSER and OCR analyses for different text patterns are described in experimental results. MSER indicates that it is a robust algorithm for text detection. Therefore, this paper analyses the detection and perception of various text patterns on different objects.

The World Health Organisation (WHO) has described the number of visually impaired people worldwide as 285 million [25]. Thirty-nine million of them are entirely blind. One of the challenging tasks that should be carried out by visually challenged is the identification of challenges that could be carried out through machine learning (ML). The Artificial Intelligence (AI) solution provides the framework with the opportunity to naturally learn and draw on knowledge without complex programming. It provides the machine with computer vision that takes judgments based on training algorithms. The main purpose of this thesis is to create an object recognition device to allow fully blind individuals to independently control their tasks. Paper often contrasts numerous target tracking algorithms, such as the Haar cascade and the CNN (CNN). The Haar Cascade classification is an important face-detection algorithm that can also be learned to recognise various objects, whereas the convolutionary neural network is highly informed and can be used for object recognition. The custom dataset consists of 2300 photographs of three groups. This analogy is rendered to find the CNN as a fitting algorithm for this method from the real-time scenario component of accuracy.

Blindness in many disorders is very widespread and lasting [26]. The World Health Organisation (WHO) reports that 285 million individuals are visually impaired. The proposed VI (Visually Disabled) Assistant Device is built to support visually impaired individuals with four modules that identify challenges, prevent hindrance, traverse both indoor and outdoor and share positions in real-time. The device suggested is a combination of smart gloves and smart phones that works well even at low light stage? The smart glove as part of the approach is used to sense and remove hazards and to recognise the environments surrounding the visually-impaired. In the area, the smartphone-based obstacle and item detectors are used to detect different objects. The device also offers smooth indoor navigation with Wi-Fi connection points enabled. The device also presents protection for the blind by the sharing of an outdoor position in real time. Our system proposed is efficient, effective, realistic and workable.

Vision is one of the key senses allowing citizens to communicate with the natural world [27]. There are about two hundred million blind people globally and visually disabled people obstruct numerous everyday practises. It is also really critical that blind people recognise their world and realise with which items they communicate. This project proposes an android framework for blind people to see cell phones through handheld devices. It combines diverse strategies to create a rich android framework that not only identifies artefacts in real time around visually disabled individuals, but also offers audio output to aid as easily as possible. SSD (Single Shot Detector) The algorithm is used to identify and track artefacts. This algorithm also provides almost reliable results for the real-time identification of particles and is proved quicker than other relative algorithms. In addition, the programme uses android tensorflow APIs and the TextToSpeech android API to provide audioperformance.

An assistive technology is proposed to offer automated navigation and guidance to visually disabled persons [28]. The device will have obstacle-free navigation and carry out real-time image processing to detect the obstacle. The method is intended to execute all activities in real time to collect all knowledge about challenges along the way for an optimal flow of information. The suggested navigation stick consists of a low cost barrier detector, barrier detection and target detector. The entire recovery method for visually disabled persons is made up of a heterogeneous collection of sensors and device parts, such as ultrasonic sensors, video, a single-board DSP processor, a wet floor and a charger. The machine learning model is used to identify objects, such that the consumer understands the world. The instruments suggested will be used to locate barriers, upstairs, downstairs, edges, potholes, pace breakers, small corridors, potholes, water-floors, etc. to provide the strongest directions for navigation. In the form of an audio prompt rather than a pulse, the output is given to render user-friendly and simple.

"independent living" for visually disabled individuals is the path to restoring independence and self-confidence [29]. Besides independence, visually disabled individuals memorise where the furniture is in the house. It may be confusing and potentially dangerous to shift the furniture or objects (e.g. keys, remote control, etc.) about. This article discusses a method that integrates profound learning with a camera with visually disabled items marking whether they have been relocated, such as chairs, essential everyday needs, etc. The mobile app alerts visually disabled people of the actual position and movement of all things. We aim to restore a clean, convenient and enjoyable indoor atmosphere via our scheme to visually disabled individuals.

Human vision plays a critical function in environmental understanding [30]. The word visual disability encompasses a vast spectrum and variation of vision, from blindness and vision deficiency; to low vision that even regular eye lenses or contact lenses cannot be restored to normal vision. Visually disabled tools may help them improve their lifestyle. This article introduces a multi-sensor device for indoor object recognition to

support visually disabled citizens. Object identification is carried out with statistical parameters using a captured image, which is further tested with the vector machine support algorithm. To boost the precision of the target detection, the interfacing Ultrasonic Sensor uses the multi-sensor principle. In addition, an infrared sensor is used to locate tiny object close foot. Experimental findings indicate the feasibility of the approach suggested. Obstacle detection and alert technology is intended to enable people with a visual disability to mitigate damage [20]. We research the obstacle detection module for stopping visually disabled people from crashing the obstructions, prepare, evaluate and improve it in depth and give it a warning message to stop an accident. Improving the standard of living to keep them comfortable and unconcerned. In terms of methodologies, a real-time camera or VDO processing device called "You Look Only Once" or YOLO is chosen. In this growth, there are two versions: YOLOv2 and YOLO9000. YOLOv2 is higher in 20:7 frames per second (FPS) than YOLO9000 and more precise. We then merge the identification and notification modules. Finally, this device is able to automatically identify and give the alarm to the visually disabled before the obstructions crash. Ception is one of the most essential senses of all current human senses, and it plays a vital function in the perception of the world [31]. It is dangerous for visually disabled individuals to step outside without any supervision. This paper is also an effort to create an object recognition method for visually disabled persons. To do this, a few segments such as a camera, an application, and an audio system are needed. We also developed and introduced an Android programme to track items surrounding the visually disabled person with the phone's camera. In addition, the programme tells the consumer of the object position and the user's distance from the object. The Application would advise the visually impaired consumer with an auditory device such as a headset or telephone speaker on the name, location and distance of the object. This device can support the visually disabled by educating them of the many things surrounding them and letting them move independently. Our aim is therefore to build a method of visual replacement that will support the visually disabled in their everyday lives by an object recognition system through which we notify them of the different items surrounding them.

Detection of moving items in real time and a step against visually disabled individuals are a difficult field of study [32]. The recent technological development for real-world capture and mobile devices such as Microsoft Kinect demands easy, accurate and faster technology to assist blind navigation. This paper aims to establish a suitable and efficient technique to transfer object detection and move indoor course. Deep details about a blind person's front scene is recorded using Microsoft Kinect. Three consecutive depth frames are extracted in one second from a film, and four-line profile lines of each depth frame are created from a distance along the line profile graph. These line profile graphs are then evaluated for the presence and trajectory of moving entity. Following an analysis of the data, the experimental findings suggest that the proposed approach can detect moving target with 92 percent precision and still objects with 87 percent precision. The average precision of the approach suggested is 90%.

Many assistance mechanisms for the identification of items (or obstacles) of significance for visually disabled persons (VIPs) have long been studied. However, the functional frameworks in the modern world are also very demanding because of standardised entity types in incredibly humiliating scenes (or complicated background) [33]. In this essay, we suggest a novel framework to detect and approximate the complete model of the typical artefacts in the everyday life of the VIP. Not only can the proposed method include relevant facts, such as size and protection instructions for catching on a flat surface, solve the issue of where the item is questioned. The pipelines use a rigorous estimator to incorporate a sequence of point cloud representation, table plane identification, entity detection and the whole model estimation. In this paradigm work, we discuss the new benefits of profound learning (e.g. RCNN, YOLO), which may be an effective means of calculating the mission when geometry-based methods approximate a complete 3D model. This scheme would not involve the isolation (or segmentation) of the items involved from the context of the scenes. The method suggested is contrasted to other methods and tested on the existing datasets gathered in typical scenes such as the kitchen or cafeteria room. The suggested structures fulfil the criteria of high precision, cycle time and suitability for VIPs in these assessments. The assessment data sets are rewritten.

Deep neural network recognition (DCNN) is an effective solution to visual perceptions that demands tremendous computational and coordination costs [34]. We have proposed a fast and low-power object recognition processor that lets visually disabled people recognise their environment. A DCNN automatic measuring algorithm has been developed that successfully quantizes data with 32 values at 8-bit fixed points and uses 5-bit indexes for display, reducing the hardware costs to marginal precision relative to the 16-bit DCNN. The design of a certain hardware accelerator uses reconfigurable process engines for multi-layer pipelines to minimise or significantly eliminate the temporary transfer of off-chip files. To significantly minimise power, a search table is used to force all multiplication in convolutions. The design is constructed of SMIC 55-nm and only 68-mw with a 1,1-v voltage of 155GB/s with a maximum output of 2.2 Top/w after a structure simulation has been done.

In this article, we established an indoor image processing method for the colour recognition of related objects [35]. Our device utilises colour recognition technologies to assess the position of a consumer with very high

precision for real-time applications. We used our method to filter a picture for a particular colour and to collect pixel coordinates for the image.

The position of the consumer is then calculated by contrasting the matrix with the pre-created matrix of the training images with these values. We have successfully performed indoor tests and have obtained really positive results. After reviewing the findings, we suggest that our localization systems be combined with indoor navigation systems, where precision is the most essential aspect for blind persons. We have also created an Android-based framework to ease the navigation phase.

III THEORETICAL FRAMEWORK

For the development of this project has been necessary to conceptualize several terms that will help to better understand the purpose of this research. These being explained from the most general aspects to the most specific concepts that must be needed for the realization of this Paper. sciences focused and applied mainly to the study of the storage, transformation and transfer of information in computers [36]. Computer science can be divided into two perspectives; a theoretical part where the design of algorithms is understood following mathematical techniques such as optimization, probabilistic techniques. The practical part will then be the implementation of these algorithms that together to form a software that will operate on a specific hardware.

Machine learning Also called machine learning or machine learning, it is a technique derived from computer science where the main objective is to make computer equipment capable of learning. In this context, learning refers to identifying patterns in millions of data and through them predicting future behaviors in an environment or situation [37] using statistical algorithms and probability theories.

Two main subareas can be identified in machine learning; supervised learning and unsupervised learning.

Through the first method, it is sought by means of the meaningful collection of a set of examples of which the answer is known, to generate a descriptive equation of the system (hypothesis) to give a possible solution to a new entry.

Two types of problems can be recognized from this same method:

- Regression problems: Its purpose is to generate a continuous value according to a series of examples with their respective answers, for example, the value of a house taking into account its area, or its location

On the other hand, unsupervised learning commonly works with random inputs to the system and the output represents the degree of familiarity or similarity between the information that is being presented to the input and the information that has been shown until then

Artificial neural networks: Neural networks have become relevant in machine learning applications due to their ability to solve non-linear equations and their cognitive learning ability [14]. Its purpose is to emulate the behavior of the nervous system and the way it can process information through neurons in conjunction with the brain.

Thus, in this paradigm of artificial intelligence there is a unit analogous to the biological neuron; the perceptron [38]. This being an element with several inputs that emulate the dendrites of the neuron and an output variable representing the

axon, the inputs that are multiplied by a vector of weights, are combined with a basic sum to later be compared or analyzed by a function that determine the output. As can be seen in Figure 1

An artificial neural network will then be a set of elementary units (Perceptron) connected in a concrete way [39].

The architecture of these neural networks is basically made up of groups of perceptrons grouped by levels or layers; an input layer where the data is entered into the network, a hidden layer and a final layer where the network outputs are represented. The layers that lie between the two mentioned above are commonly known as hidden layers.

Vector support machines: Support Vector Machines (SVM) bring together a series of algorithms developed in the 90s by Vladimir Vapnik. Initially the support machines were conceived to solve binary classification problems, over time they were also lending themselves to develop regression and multiclass classification problems.

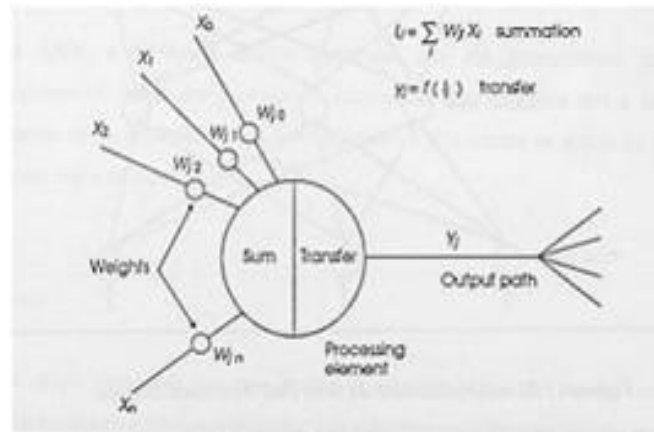


Fig. 1. Basic perception Model

The main objective of the SVM's is to build a hyperplane that functions as a dividing line between examples of different classes, in a similar way as the limit boundaries that make artificial neural networks (decision boundaries) do. The biggest difference between the hyperplanes that the VM's and the borders of traditional neural networks, is that SVM's seek to maximize the margin between classes, that is, the decision border must be at the greatest possible distance from the classes that it is separating.

The hyperplane is a one-dimensional subspace described by the following equation:

$b + W_1 * X_1 + W_2 * X_2 + W_3 * X_3 + \dots = 0$ (1). Or as can be seen in figure 2; simply a vector of weights (W) transposed

by a vector of characteristics (X).

reference to each of the classes. This is where the concept of reliability arises, the more positive the value, the more reliable it will be of one class and the more negative, the more reliability it will have of the other, as can be seen in figure3

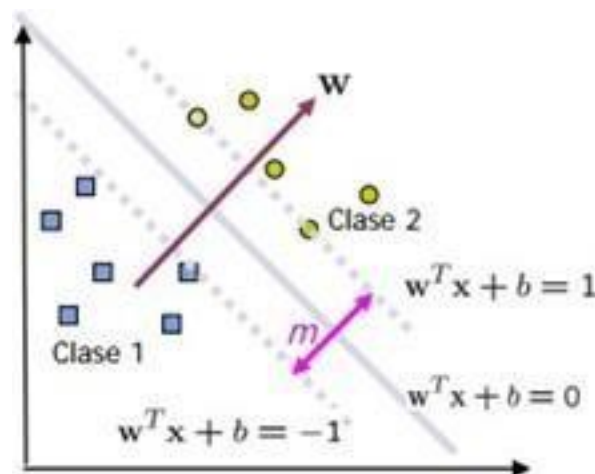


Fig. 2. Example of the hyperplane built by an SVM to separate two classes, image taken VECTOR SUPPORT MACHINES, Gustavo A. Betancourt

In the same way as can be seen in figure 2, the value of the hyperplane can take positive or negative values, making The maximum margin hyperplane is the desired optimal subspace, this depends on determining the maximum distance that exists between the distances that are created with all possible hyperplanes and the points closest to it. These points are called vector support points.

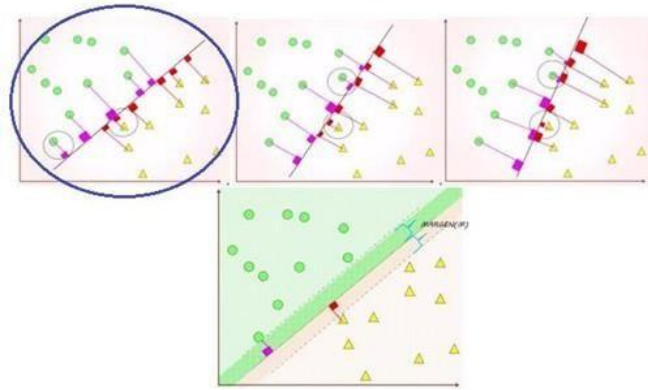


Fig. 3 Image exemplifying the determination of the best hyperplane,taken from the video "Vector support machine"

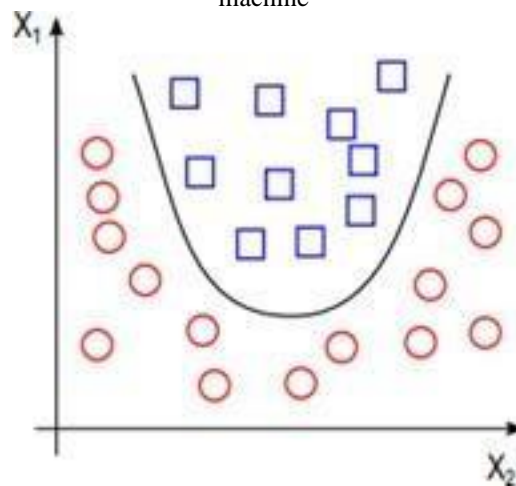


Fig. 5. Representative figure of a non-linearly separable casetaken Tutorial on SupportVector Machines, Enrique Carmona

One of the most important components in the construction of SVM's is the Kernel to be used. Support machines are commonly known as linear separators, but there are cases where a straight line is not enough to make the separation between categories. Cases such as, for example, where the data separation cannot be done completely, or where it is necessary to classify more than two categories. In these situations, it is necessary to make non-linear separation curves. Adding Kernel allows to make the built hyperplane more flexible.

Deep learning: Deep learning is a subset of machine learning. It is used for problems where traditional learning methods do not achieve an appropriate performance, deep learning is inspired by the brain to create neural networks with a large number of hidden layers compared to traditional neural networks. This type of topology allows these "deep neural networks" to obtain simple patterns or characteristics from complex inputs.

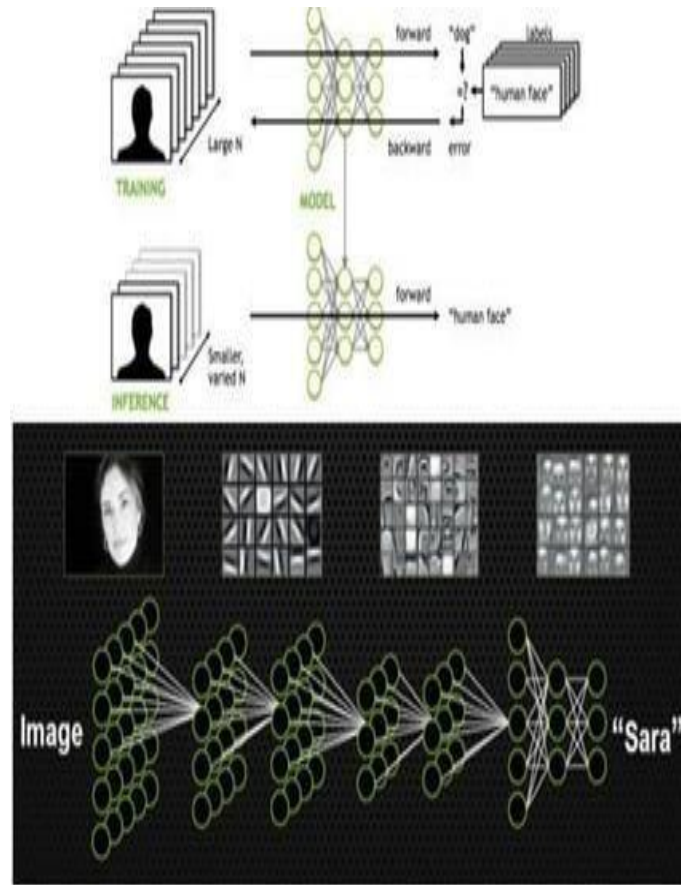
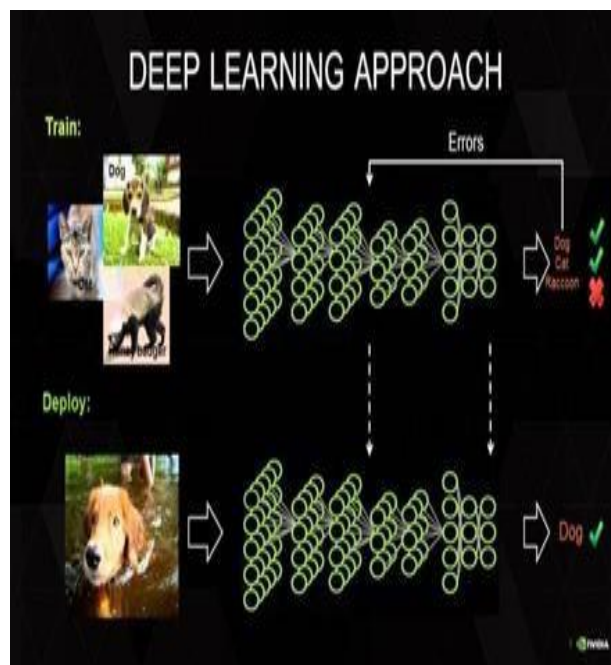


Fig. 6. Representative image of the operation of Deep learning for the identification of faces, Image taken from <https://www.quora.com/What-is-deep-learning>

Deep learning techniques are not only used for classification problems, they also perform well in unsupervised learning problems, such as pattern recognition. On the other hand, speech and text processing are also areas of interest for researchers of these promising techniques.



Web services: Web services allow communication between applications that were developed in different languages and even that work with different operating systems, web services need a well-structured architecture to allow such interaction on the web, they have well-defined protocols for the exchange of information such as SOAP, WS-addressing and MTOM, which are basically in charge of exchanging information in an optimal way. Another fundamental component in web services are the service descriptors known as WSDL and WS-CDL, which are in charge of describing the services as such that WEB services should do. Today they are widely used for their versatility and are generally used by mobile applications.

IV CONCLUSION

The design of multiple One class vector support machines as a classification system proved to be very efficient for the development of this project, as the fact of associating an SVM to each of the categories gave great versatility. In this way, the training of each one of them (both new and existing) can be carried out independently without depending on the rest of the system. In addition to the fact that this architecture allows each class to grow at its own pace and that the quality of the classifier is not affected by an imbalance in terms of the number of examples for each of the categories (a situation that would affect in case of choosing another type of architecture such as One Vs All). An automatic grid search is a key factor when training each of the classes in the system. As the optimal value of the parameters of the SVMs used changes from class to class. This added to the design of multiple SVMs (one per class) allows the system to become completely autonomous and to be able to continuously improve thanks to the data provided by the users. This continuous improvement refers not only to the increased accuracy in existing classes as your training data grows, but also the ability to learn new classes supplied by users. Initially, it was sought to design a classification methodology based on the similarity and context of the objects. Throughout the development of the project, it was possible to conclude that the inclusion of AlexNet's pre-trained network would allow to generalize all the objects that make up the created database. This is due to the fact that the aforementioned network was trained with 1 million images distributed among 1000 different categories, allowing to avoid the grouping of objects according to similarity and context. For this reason, it was also determined to change the implementation of a set of applications that allow user-server interaction for a single application capable of recognizing any type of object.

REFERENCES

1. C. Park, S. W. Cho, N. R. Baek, J. Choi, and K. R. Park, "Deep Feature-Based Three-Stage Detection of Banknotes and Coins for Assisting Visually Impaired People," *IEEE Access*, vol. 8, pp. 184 598– 184 613, 2020. [Online]. Available: [10.1109/access.2020.3029526](https://doi.org/10.1109/access.2020.3029526); <https://dx.doi.org/10.1109/access.2020.3029526>
2. S. Bhatlawande, M. Mahadevappa, J. Mukherjee, M. Biswas, D. Das, and S. Gupta, "Design, Development, and Clinical Evaluation of the Electronic Mobility Cane for Vision Rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 6, pp. 1148–

- 1159,2014.[Online].Available: 10.1109/tnsre.2014.2324974;https://dx.doi.org/10.1109/tnsre.2014.2324974
3. Aladren, G. Lopez-Nicolas, L. Puig, and J. J. Guerrero, "Navigation Assistance for the Visually Impaired Using RGB-D Sensor with Range Expansion," *IEEE Systems Journal*, vol. 10, no. 3, pp. 922–932, 2016. [Online]. Available: 10.1109/jsyst.2014.2320639; https://dx.doi.org/10.1109/jsyst.2014.2320639
 4. D. Dakopoulos and N. G. Bourbakis, "Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 1, pp. 25–35, 2010. [Online]. Available: 10.1109/tsmcc.009.2021255;https://dx.doi.org/10.1109/tsmcc.2009.2021255
 5. Ton, "LIDAR Assist Spatial Sensing for the Visually Impaired and Performance Analysis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, pp. 1727–1734, 2018.
 6. M.-C. Kang, S.-H. Chae, J.-Y. Sun, J.-W. Yoo, and S.-J.Ko, "A novel obstacle detection method based on deformable grid for the visually impaired," *IEEE Transactions on Consumer Electronics*, vol. 61 no. 3, pp. 376–383, 2015. [Online]. Available: 10.1109/tce.2015.7298298; https://dx.doi.org/10.1109/tce.2015.7298298
 7. E. Cardillo, "An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking," *IEEE Sensors Journal*, vol. 18, no. 6, pp. 2568–2576,2018.
 8. J. Chen, Z. Xu, and Yu, "A 68-mw 2.2 Tops/w Low Bit Width and Multiplierless DCNN Object Detection Processor for Visually Impaired People," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, pp. 3444–3453, 2019.
 9. B. Ando and S. Graziani, "Multisensor Strategies to Assist Blind People: A Clear-Path Indicator," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 8, pp. 2488–2494, 2009. [Online]. Available: 10.1109/tim.2009.2014616;https://dx.doi.org/10.1109/tim.2009.2014616
 10. S. Pehlivan, M. Unay, and A. Akan, "Designing an Obstacle Detection and Alerting System for Visually Impaired People on Sidewalks," in *2019 Medical Technologies Congress (TIPTEKNO)*, 2019, pp. 1–4.
 11. F. Al-Muqbali, N. Al-Tourshi, K. Al-Kiyumi, and F. Hajmohideen, "Smart Technologies for Visually Impaired: Assisting and conquering infirmity of blind people using AI Technologies," in *2020 12th Annual Undergraduate Research Conference on Applied Computing (URC)*, 2020, pp.1–4.
 12. A. N. Zereen and S. Corraya, "Detecting real time object along with the moving direction for visually impaired people," in *2016 2nd International Conference on Electrical*, 2016, pp.1–4.
 13. W. Lin, M. Su, W. Cheng, and W. Cheng, "An Assist System for Visually Impaired at Indoor Residential Environment using Faster-RCNN," in *2019 8th International Congress on Advanced Applied Informatics (IIAI- AAI)*, 2019, pp. 1071–1072.
 14. S. Deshpande and R. Shriram, "Real time text detection and recognition on hand held objects to assist blind people," in *2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT)*, 2016, pp.1020–1024.
 15. N. Khaled, S. Mohsen, K. E. El-Din, S. Akram, H. Metawie, and A. Mohamed, "In-Door Assistant Mobile Application Using CNN and TensorFlow," *2020 International Conference on Electrical*.
 16. S. Shah, J. Bandariya, G. Jain, M. Ghevariya, and S. Dastoor, "CNN based Auto-Assistance System as a Boon for Directing Visually Impaired Person," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 235–240.
 17. T. Saeteng, T. Srionuan, C. Choksuchat, and N. Trakulmaykee, "Reforming Warning and Obstacle Detection Assisting Visually Impaired People on mHealth," in *2019 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia)*, 2019, pp. 176–179.
 18. K. Shishir, S. R. Fahim, F. M. Habib, and T. Farah, "Eye Assistant: Using mobile application to help the visually impaired," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 2019, pp.1–4.
 19. K. Lakde and P. S. Prasad, "Navigation system for visually impaired people," in *2015 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)*, 2015, pp. 93– 0098.
 20. A. Adishesha and B. Desai, "3D Imprinting of the Environment for the Visually Impaired," in *2015 IEEE European Modelling Symposium (EMS)*, 2015, pp. 148–153.
 21. S. Yadav, R. C. Joshi, M. K. Dutta, M. Kiac, and P. Sikora, "Fusion of Object Recognition and Obstacle Detection approach for Assisting Visually Challenged Person," in *2020 43rd International Conference on Telecommunications and Signal Processing*, 2020, pp. 537–540.
 22. S. Alghamdi, R. V. Schyndel, and I. Khalil, "Safe trajectory estimation at a pedestrian crossing to assist visually impaired people," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 5114–5117.

23. T. Patel, V. J. Mistry, L. S. Desai, and Y. K. Meghrajani, "Multisensor - Based Object Detection in Indoor Environment for Visually Impaired People," in 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018, pp. 1–4.
24. M. Kamal, A. I. Bayazid, M. S. Sadi, M. M. Islam, and N. Hasan, "Towards developing walking assistants for the visually impaired people," in 2017 IEEE Region 10 Humanitarian Technology Conference (R10- HTC), 2017, pp.238–241.
25. H. Dahiya, M. K. Gupta, and Dutta, "A Deep Learning based Real Time Assistive Framework for Visually Impaired," in 2020 International Conference on Contemporary Computing and Applications (IC3A), 2020, pp.106–109.
26. Adishesha and B. Desai, "3D Imprinting of the Environment for the Visually Impaired," *2015 IEEE European Modelling Symposium (EMS)*, Madrid, Spain, 2015, pp. 148-153.doi: 10.1109/EMS.2015.32
27. C. K. Lakde and P. S. Prasad, "Navigation system for visually impaired people,"*2015 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)*, Melmaruvathur, India, 2015, pp. 0093-0098.doi: 10.1109/ICCPEIC.2015.7259447
28. S. Alghamdi, R. van Schyndel and I. Khalil, "Safe trajectory estimation at a pedestrian crossing to assist visually impaired people," *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, USA, 2012, pp. 5114- 5117.doi: 10.1109/EMBC.2012.6347144
29. Hangrong Pan, C. Yi and Y. Tian, "A primary travelling assistant system of bus detection and recognition for visually impaired people," *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, San Jose, CA, USA, 2013, pp. 1- 6.doi:10.1109/ICMEW.2013.6618346
30. R. Tapu, B. Mocanu, A. Bursuc and T. Zaharia, "A Smartphone-Based Obstacle Detection and Classification System for Assisting Visually Impaired People," *2013 IEEE International Conference on Computer Vision Workshops*, Sydney, NSW, Australia, 2013, pp. 444- 451.doi: 10.1109/ICCVW.2013.65
31. N. Mahmud, R. K. Saha, R. B. Zafar, M. B. H. Bhuiyan and S. S. Sarwar, "Vibration and voice operated navigation system for visually impaired person," *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, Dhaka, Bangladesh, 2014, pp. 1-5.doi: 10.1109/ICIEV.2014.6850740
32. Noorithaya, M. K. Kumar and A. Sreedevi, "Voice assisted navigation system for the blind," *International Conference on Circuits, Communication, Control and Computing*, Bangalore, India, 2014, pp. 177-181.doi: 10.1109/CIMCA.2014.7057785:
33. S. Abdullah, N. M. Noor and M. Z. Ghazali, "Mobility recognition system for the visually impaired," *2014 IEEE 2nd International Symposium on Telecommunication Technologies (ISTT)*, Langkawi, Malaysia, 2014, pp. 362-367.doi: 0.1109/ISTT.2014.7238236
34. S. Chae, J. Sun, M. Kang, B. Son and S. Ko, "Collision detection based on scale change of image segments for the visually impaired," *2015 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 2015, pp. 511-512.doi: 10.1109/ICCE.2015.7066504
35. M. Nie *et al.*, "SoundView: An auditory guidance system based on environment understanding for the visually impaired people," *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN, USA, 2009, pp. 7240- 7243.doi:10.1109/IEMBS.2009.5334754
36. S. F. Toha, H. M. Yusof, M. F. Razali and A. H. A. Halim, "Intelligent path guidance robot for blind person assistance," *2015 International Conference on Informatics, Electronics & Vision (ICIEV)*, Fukuoka, Japan, 2015, pp. 1-5.doi: 10.1109/ICIEV.2015.7334040
37. R. Bostelman, P. Russo, J. Albus, T. Hong and R. Madhavan, "Applications of a 3D Range Camera Towards Healthcare Mobility Aids," *2006 IEEE International Conference on Networking, Sensing and Control*, Ft. Lauderdale, FL, USA, 2006, pp. 416- 421.doi: 10.1109/ICNSC.2006.1673182
38. H. Harms, E. Rehder, T. Schwarze and M. Lauer, "Detection of ascending stairs using stereo vision," *2015IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Hamburg, Germany, 2015, pp. 2496- 2502.doi: 10.1109/IROS.2015.7353716
39. M. Vlaminck, L. Jovanov, P. Van Hese, B. Goossens, W. Philips and A. Pižurica, "Obstacle detection for pedestrians with a visual impairment based on 3D imaging," *2013 International Conference on 3D Imaging*, Liege, Belgium, 2013, pp. 1-7.doi: 10.1109/IC3D.2013.6732091