Journal of Information Technology Education: Research

An Official Publication of the Informing Science Institute
InformingScience.org

JITEResearch.org

Volume 22, 2023

# A New Knowledge Primitive of Digits Recognition for NAO Robot Using MNIST Dataset and CNN Algorithm for Children's Visual Learning Enhancement

| | | |
|---|---|---|
| Soukaina Gouraguine* | ENSET of Mohammedia, Hassan II University, Casablanca, Morocco | soukainagoura@gmail.com |
| Mohammed Qbadou | ENSET of Mohammedia, Hassan II University, Casablanca, Morocco | qbmedn7@gmail.com |
| Mohamed Rafik | ENSET of Mohammedia, Hassan II University, Casablanca, Morocco | rafik@enset-media.ac.ma |
| Mustapha Riad | ENSET of Mohammedia, Hassan II University, Casablanca, Morocco | My.mustapha.riad@gmail.com |
| Khalifa Mansouri | ENSET of Mohammedia, Hassan II University, Casablanca, Morocco | khalifa.mansouri@enset-media.ac.ma |

* Corresponding author

## Abstract

| | |
|---|---|
| Aim/Purpose | Our study is focused on prototyping, development, testing, and deployment of a new knowledge primitive for the humanoid robot assistant NAO, in order to enhance student visual learning by establishing a human-robot interaction. |
| Background | This new primitive, utilizing a convolutional neural network (CNN), enables real-time recognition of handwritten digits captured by the NAO robot, a humanoid robot assistant developed by SoftBank Robotics. It is equipped with advanced capabilities, including a wide range of sensors, cameras, and interactive features. By integrating the proposed primitive, the NAO robot gains the ability |

Accepting Editor Janice Whatley │Received: May 12, 2023│ Revised: July 24, August 12, September 11, 2023 │ Accepted: September 26, 2023.
Cite as: Gouraguine, S., Qbadou, M., Rafik, M., Riad, M., & Mansouri, K. (2023). A new knowledge primitive of digits recognition for NAO robot using MNIST dataset and CNN algorithm for children's visual learning enhancement. *Journal of Information Technology Education: Research, 22,* 389-408. https://doi.org/10.28945/5194

| | |
|---|---|
| | to accurately recognize handwritten digits, contributing to improved student visual learning experiences. |
| Methodology | Our developed primitive consists of the use of a convolutional neural network (CNN) so that the robot is able to recognize the handwriting of the digits present in the input image received in real-time. The NAO robot establishes interaction with the learners through a scenario based on a predefined assignment. In this scenario, NAO captures the digit handwritten by the learner via its camera, recognizes the digit using the deep learning model generated by the MNIST dataset, and announces to the learner the handwritten digit in the input image. The prototype is realized using the concept of a distributed system allowing the distribution of tasks in four different computing nodes. |
| Contribution | Our research makes a significant contribution by equipping the humanoid robot NAO with a cognitive intelligence system through the integration of a new knowledge primitive based on handwriting digit recognition (HWDR). Our approach used to create and implement this primitive in the NAO robot is interesting and innovative, and presents a promising provision for enhancing the visual learning experience of children and young students with special needs, based on the use of distributed systems that divide the work using various components distributed over several nodes, coordinating their efforts to perform tasks more efficiently than a single device besides the NAO robot. |
| Findings | We designed our model using specific parameters and a fully convolutional neural network architecture, which includes three residual depthwise separable convolutions, each followed by batch normalization and ReLU activation. To evaluate the performance of our model, we tested it on the MNIST dataset, where we achieved a remarkable accuracy, F1 score, and recall of 99%. An experiment was conducted to test our implemented primitive and see the effectiveness of this invention for enhancing visual learning in children with special needs. We developed a visual learning strategy based on the creation of engaging activities mediated by the NAO robot in an educational context. The results showed that participants achieved a strong commitment to the NAO robot, appreciating its ability to recognize handwritten digits and highlighting its promising potential to enrich visual learning experiences. Participants expressed a strong preference for teaching methods integrating assistive learning technologies, demonstrating the positive impact of our humanoid assistant robot on improving learning and visual intelligence in an educational environment. |
| Recommendations for Practitioners | Encourage creativity and innovation in the field of robotics and special needs. This can lead to new and effective solutions that improve the lives of students with special needs. |
| Recommendations for Researchers | Test and evaluate the proposed robotics solutions to ensure they are effective and making a positive impact. Use feedback from users, educators, and parents to refine and improve your solutions. Also, ensure that the robotics solutions are accessible to students with a range of abilities. This may involve designing solutions that are adjustable or providing alternative means of access. |
| Impact on Society | As there are several ways to educate, there are multiple forms of learning. With the help of this learning procedure and strategy, the human teacher collaborates with the robot assistance NAO to improve visual learning among students. The findings of this research can serve as an application for the implementation of |

various pedagogical methods that will assist in meeting the needs of the majority of learners.

Future Research — Our future research will concentrate on addressing the educational needs of students with special needs, enabling them to overcome their challenges and reach academic excellence in an inclusive environment. To achieve this goal, we plan to leverage the capabilities of social robots, which have emerged as a significant contributor to the field of human-robot interaction, particularly in facilitating inclusive education. These agents have proven to be effective in providing support to students with special needs, thereby enabling them to receive the education they need to succeed.

Keywords — educational robotics, students with special needs, human-robot interaction, NAO robot, recognition of handwritten digits, convolutional neural network, visual learning

# INTRODUCTION

In the field of human-robot interaction (HRI), the incorporation of AI-enabled social robots, such as assistive technologies, is redefining the educational role by enabling instructors to be strategic supervisors of the automated learning process (Henschel et al., 2020). Educational robots collaborate with various disciplines (Woo et al., 2021), enhancing the curriculum and providing opportunities for problem-solving and innovation (Edwards et al., 2018). However, the prospect of equipping robots with cognitive intelligence to enhance authentic, intensive, and natural interaction is a challenge and source of intrigue in the field of human-robot interaction (Yuan et al., 2021). The robot's ability to understand, interpret, process, reason, and learn is key to achieving this goal (Filippini et al., 2021). The enhancement of interaction with tasks involving object recognition, image, and sound classification requires deep learning techniques, notably convolutional neural networks (CNN) which have achieved state-of-the-art results in these areas (Dhillon & Verma, 2020). These tasks require CNN architectures with varying parameters, making their deployment in robotic platforms and real-time systems challenging and often impractical. The algorithms are written in Python, and use software development tools such as TensorFlow, as well as techniques of image processing with consistent use of Python Image Library (PIL) functions, Open CV library of Open-Source Computer Vision, Keras, and others (Karayaneva & Hintea, 2018b).

Our innovative research involves developing a distributed system for real-time handwriting recognition by the humanoid robot, utilizing a high-performance CNN. This integration expands educational robots' capabilities, transforming NAO into an intelligent tool for recognizing handwritten digits, enhancing interactive visual learning experiences, and providing unique opportunities for students with special needs. With the help of this learning procedure, the human teacher collaborates with the robot assistance NAO to improve visual learning among students. The general impression is that certain students learn better when the material is presented visually. Accessible education is the process of designing a course and developing an instructional method that meets the needs of people with a diversity of life experiences, abilities, and learning styles (Dhillon & Verma, 2020). As there are several ways to educate and multiple forms of learning, the findings of this research can serve as an application for the implementation of various pedagogical methods that will help to meet the needs of the majority of learners.

The structure of this article is as follows. We start by reviewing previous work on the art of the main concepts addressed by our study. Then, we present the methodology applied, the architecture, the API, and the set of technologies used to integrate our CNN into the NAO robot. Following our comprehensive testing and evaluation, we provide a detailed explanation and schematization of the achieved results, showcasing the effectiveness of our prototype in enhancing student visual learning through human-robot interaction. Finally, we discuss the obtained outcomes.

# RELATED WORKS

## *NAO ROBOT AT SCHOOLS*

Due to the rapid and phenomenal advancements and assistive technology (AT) in all aspects of scientific, social, and professional life, dependence on robotic technology has increased, imposing its presence and importance in a variety of fields, including medical, industrial, construction, military, and civil, to the point where the industrial strength of a country is now measured by the level of development of its robots (Bouck & Long, 2021). As a result, this technology must have a strong and effective presence in the most crucial sector of any society, which is education, seeking excellence, development, and prosperity (Jung & Won, 2018).

Hence, the educational robot was one of the most prominent of these recent innovations (Belpaeme et al., 2018). NAO robots have been widely used to improve children's learning in many parts of the world (Karayaneva & Hintea, 2018b; Tamakloe, 2020), and meet the diverse needs of learners (Heinmäe et al., 2022). The NAO robot provides an educational environment that encourages teamwork and improves communication, role exchange, and decision-making skills. It has also been widely used to support the educational process and to give a secure connection to the participants concerned (teachers and learners). The findings of the study conducted by Neumann (2020) contribute to the expanding knowledge of the potential of social robots to facilitate early language and literacy skills in children. The research indicates that children engage positively with social robots through playful activities and interactions, which suggests the potential effectiveness of this approach.

## *OBJECT RECOGNITION ALGORITHMS IMPLEMENTED ON NAO ROBOT*

There is a limited amount of research on implementing object recognition algorithms on the NAO robot. For instance, while manual coding can be used for tasks such as recognizing colors, shapes, typed words, operators, and digits, a machine-learning approach using five different classifiers (Random Forest, Stochastic Gradient Descent, Support Vector Machine, Nearest Neighbors, and Neural Networks) has been developed for object classification and recognition (Karayaneva & Hintea, 2018a). This project aimed to create an object recognition application that could be tested on the NAO robot and serve as a resource for future developments in institutional settings (Karayaneva & Hintea, 2018b).

Several studies have explored the use of deep learning techniques for NAO robot detection. For example, Cruz et al. (2018) achieved a detection rate of 97% by employing two different detectors based on the XNOR-Net and Squeeze Net architectures in a robotic football setting previously used by Schnekenburger et al. (2017). These detectors enable the NAO robot to detect other robots on the field, which is crucial for team coordination and gameplay. Similarly, Albani et al. (2017) developed a deep learning approach to train the NAO robot to recognize specific objects such as a ball, lines, and goals, which are crucial for participation in the annual RoboCup competition. The use of deep learning techniques in NAO robots has shown great potential for enhancing their capabilities and expanding their use in various settings.

## *VISUAL TEACHING STRATEGIES TO IMPROVE LEARNING*

Visual learning strategies are widely considered to be effective tools for improving student learning (Munna & Kalam, 2021). These strategies use visual aids such as images, videos, graphs, and diagrams to facilitate comprehension and retention of information. Teachers have found that visual aids enable students to better visualize abstract concepts, reinforce their memory, and facilitate overall understanding of the content (Zhexenova et al., 2020). Research has shown that the use of visual teaching strategies can improve academic performance, student engagement, and motivation to learn (Al-Khresheh et al., 2020). Visual teaching strategies also offer benefits for students with special needs and learning difficulties (Atanga et al., 2020). They can be adapted to meet the specific needs of learners, using color schemes, specific images, or visual modifications to facilitate understanding. However, based on the theoretical background, we found that research has explored the use of social robots for collaborative learning to teach linguistic, social, and emotional skills to children, where they can engage in conversations, model appropriate behaviors, and provide support to students who may have difficulties in their social interactions (Desideri et al., 2018; Zeaiter, 2020).

Our proposed new approach aims to extend these efforts by exploring the potential of social robotics in the development of visual learning strategies. Our main objective is to help students with special needs and learning difficulties retain information more effectively. By exploiting the advantages of educational robots, our research offers new perspectives for the development of visual teaching strategies adapted to the needs of learners.

# METHODOLOGY

## *PROTOTYPING METHODOLOGY FOR OUR NEW PRIMITIVE ON THE NAO ROBOT*

The research work presented in this paper employs design science research (DSR) strategies (vom Brocke et al., 2020), which is a paradigm of problem-solving aimed at improving human knowledge through the creation of innovative artifacts. The stages of artifact development resemble the SDLC paradigm for software development lifestyle (Acharya & Sahu, 2020), which includes analyzing problems, planning, implementing, testing, and finally deploying the software system for real-world use. In this research, the prototyping methodology used focuses on implementing a complete system that allows the humanoid robot NAO to recognize handwritten digits using the techniques of deep learning. Figure 1 depicts the main phases of the prototyping process, which begins with the creation of a prototype – a preliminary software system version used to illustrate concepts, examine design possibilities, and get a greater knowledge of the issue and potential solutions.

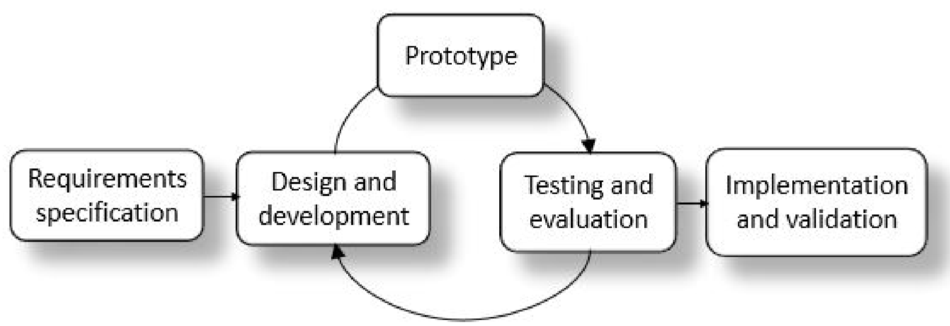



**Figure 1. Steps in prototype development for the primitive implementation**

The prototyping model begins with the collection of a requirements analysis, where the prototype's goals and scope are specified based on extensive research of existing studies. The preliminary prototype design is part of the second phase. This is a specification of the designs, frameworks, and Python libraries that have been considered, including OpenCV, PIL, TensorFlow, Keras, etc., for the development of the initial prototype. To evaluate the prototype, the algorithms' implementation entails processing the image captured by the NAO robot. Then, we focus on refining the prototype until all requirements are satisfied. Our approach to prototyping focuses on student testing, strategically integrated into the final implementation and validation phase. The tests are designed to assess the performance of the prototype and its ability to assimilate intellectual knowledge for the humanoid robot. This integral step reflects our commitment to creating a prototype that not only meets technical criteria but is also optimized for effective engagement and pedagogical value. As part of this process, student interactions play a key role in refining the prototype's design and functionality, reinforcing its alignment with our educational goals. The following section elaborates on the detailed progression and deployment of the fundamental module enabling the NAO robot to recognize handwritten characters, in particular numbers.

## *HANDWRITTEN RECOGNITION PROCESS WITH NAO ROBOT*

Following the steps of prototype development, in this section, we propose an efficient flowchart describing the experimental procedure for the prototype of the detection and recognition of handwritten digits by users from the NAO humanoid robot. First, the program allows access to the real-time feed of the visual input (webcam) of the NAO robot, and then the captured frame containing the handwritten digit is directly fed to the CNN. Then a CNN-based classification is performed, allowing the communication of the output class to the NAO robot so that the NAO robot can recognize the handwritten digit appropriately. The flowchart of the primitive behavior programming for the NAO robot is shown in Figure 2.

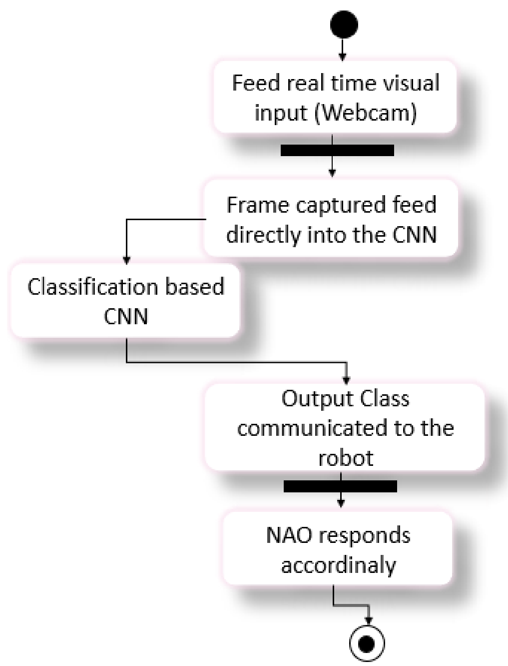



**Figure 2. Flowchart of the handwritten recognition primitive for NAO**

## *SYSTEM DESIGN OF THE PROTOTYPE*

Our research uses a human-robot interaction-based scenario in which the NAO robot interacts with learners to simulate the process of recognizing handwritten digits. The only element and entity of the prototype system with which the user interacts is the assistant robot NAO. NAO serves as the user's front-end interface to the prototype system. To establish the interaction, a number of the robot's components and sensors are available. Figure 3 shows some of the critical modules used in the development of this prototype as well as the technical architecture model used to create our prototype.

The second part, representing a distributed system, essentially serves as the prototype's back-end system. The primary server in charge of communicating with the humanoid robot NAO and the HWDR divides the implementation task of the prototype into several parts, and several nodes to achieve a correct prediction that can be announced by the robot. This could be considered an advantageous, efficient, and powerful solution if several robots are sending multiple images at the same time.

Thus, when a task workload is too large to be managed by a single computer or machine, distributed systems are useful in these situations to help improve performance by providing each node with the ability to process different parts of a task simultaneously, while ensuring scalability, resiliency, redundancy, cost-effectiveness, and efficiency. The main server is primarily responsible for receiving the image captured by the robot, pre-processing the image for model compatibility, to be consistent with the model, loading and executing the predictions through the entrained model, performing additional operations on the predicted results, and then finally sending the results to the NAO. The central server is essentially responsible for creating an API for deep learning models. It must deal with the deep learning frameworks that the HWDR model was trained on. The necessary frameworks are TensorFlow, Keras, OpenCV, and the Pillow libraries.

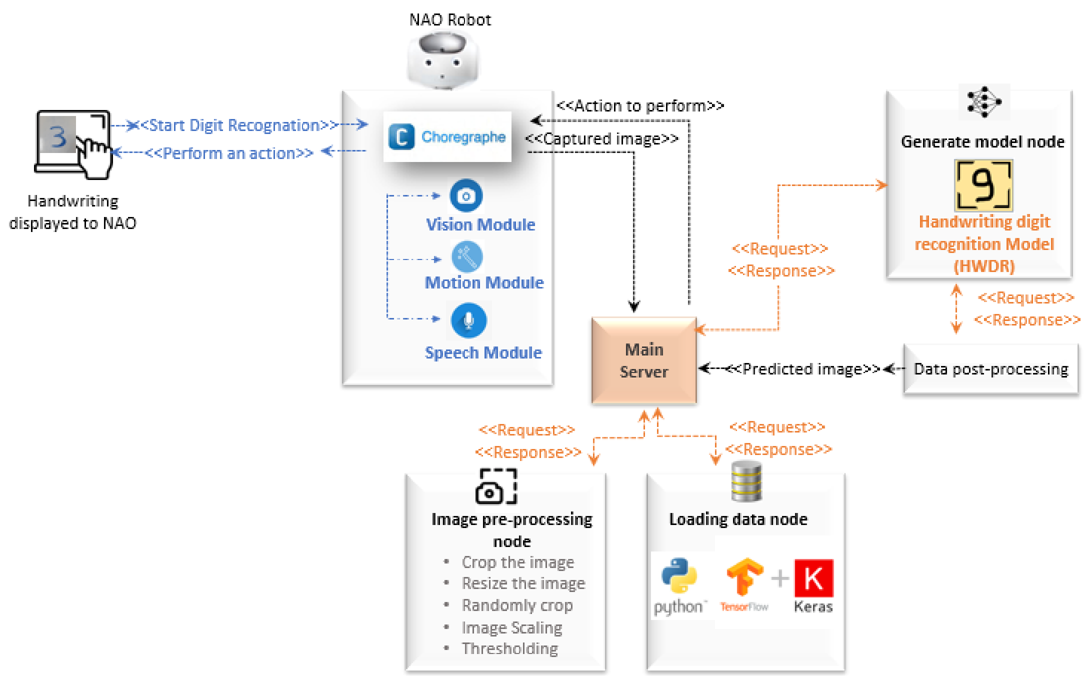



**Figure 3. Technical architecture model**

The NAO robot offers a Python development kit called NAOqi, which enables researchers to create advanced intelligent components for various robotic applications, including motion processing, speech recognition, and vision processing. NAOqi is the software framework that powers the NAO robot (Alam et al., 2016). NAOqi is a C++ and Python cross-platform programming framework. The Choregraphe programming tool, which is recommended by robot manufacturers, is a better option for programming and creating activities for the NAO robot (Pot et al., 2009). Choregraphe is a desktop application that runs on multiple platforms and allows the creation of complex scenarios, animations, and behaviors, such as user interaction, custom response animations, and voice recognition.

## *Convolutional Neural Network Implementation*

Our research focuses on using deep learning (DL) methods to recognize handwritten digits in an input image. As with any deep learning model, it requires a substantial amount of high-quality training data to achieve superior performance (Aly & Dugan, 2017). Therefore, we have provided a quantitative and qualitative description of the dataset used to train the deep learning model in the subsequent section. Furthermore, we have provided an overview of the key parameters utilized in developing our handwritten character recognition model with NAO (HWDR). This will help readers to better understand the dataset and parameters used to achieve a high level of accuracy in our model.

## MNIST database of handwritten digits

The MNIST Mixed National Institute of Standards and Technology dataset is a widely-used image database utilized by numerous image processing systems and machine learning applications (Cohen et al., 2017). It consists of 70,000 grayscale images, each with a size of 28x28 pixels. The dataset is further divided into 60,000 training images and 10,000 test images, representing handwritten digits from zero to nine. The exact number of images for each digit can be found in Table 1. Due to its ubiquity, the MNIST dataset has become a standard benchmark dataset for evaluating the accuracy of image recognition algorithms and is commonly used in both research and educational settings.

**Table 1. Number of images for each class**

| CLASSES | TRAINING | TEST | TOTAL |
|---|---|---|---|
| 0 | 5923 | 980 | 6903 |
| 1 | 6742 | 1135 | 7877 |
| 2 | 5958 | 1032 | 6990 |
| 3 | 6131 | 1010 | 7141 |
| 4 | 5842 | 982 | 6824 |
| 5 | 5421 | 892 | 6313 |
| 6 | 5918 | 958 | 6876 |
| 7 | 6265 | 1028 | 7293 |
| 8 | 5851 | 974 | 6825 |
| 9 | 5949 | 1009 | 6958 |
| Total | 60000 | 10000 | 70000 |

The MNIST dataset thus consists of 10 different classes, ranging from zero to nine. The handwritten digit images are presented as a 28×28 matrix where each cell consists of a grayscale pixel value, a qualitative description is shown in Figure 4 allowing us to visualize the clarity of the images and their allocation to the appropriate classes.

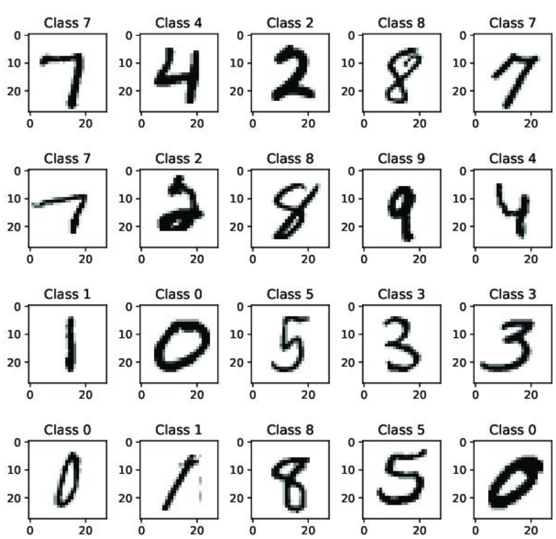



**Figure 4. Examples of digits**

### Deep learning architecture for HWDR

Table 2 illustrates the CNN architecture for measuring and establishing the performance score for HWDR as well as the parameters for configuration. The seven layers include convolutional layers, activation functions, and max-pooling layers, which allow for the extraction of features and patterns from the input images. The choice of hyperparameters, such as the learning rate, batch size, and number of epochs, can significantly affect the performance of our CNN model.

The CNN architecture was applied to the HWDR dataset for handwritten digit recognition. By using a training batch size of 32 and a learning rate of 0.0001, the CNN was trained for 50 epochs. The RelU activation function was applied after each convolution layer to introduce non-linearity, and the max-pooling layer was used to reduce the size of the feature maps and remove noise. The resulting high training score and good test score indicate that the CNN architecture and hyperparameters were well-suited for the HWDR dataset and demonstrate the effectiveness of deep learning approaches for image recognition tasks.

**Table 2. The architecture of our CNN**

| PARAMETERS | BASELINE |
|---|---|
| Batch_size | 32 |
| Number of layers | 7 Layers |
| Number of epochs | 50 |
| Number of steps | 32 |
| Learning rate | 0.0001 |
| Activation function | ReLU |

### *Participants in Our Experimental Study*

We tested and validated our approach by creating an innovative educational environment and developing interactive scenarios mediated by our assistant robot, NAO, for students with special needs. We have recruited a group of 12 Moroccan children with special needs, aged between 5 and 7, to take part in our experiment, which was carried out in the Laboratory for Modelling and Simulation of Intelligent Industrial Systems located at the Higher Normal School of Technical Education of Mohammedia (ENSET), an engineering school in Morocco. The group was equally divided between the sexes, with six girls and six boys (Average age = 6 years, St Dev age = 0.8 years). By focusing on a balanced and representative sample of children with specific needs, we aimed to provide robust and meaningful results to support our conclusions. Participants were carefully recruited from a variety of sources, to ensure that our study reflects a diverse and representative population of the children our approach seeks to benefit. This experimental approach allows us to better understand the impact of our visual learning strategy on children with special needs and to help improve their visual learning experiences.

## Stages of Prototyping And Evaluation

### *The Architecture Model Resulting From Our CNN*

The model architecture of the convolutional neural network (CNN) conceived and used for handwritten digit recognition (HWDR) in our investigation is shown in Figure 5. The initial image captured by the NAO robot is preprocessed into a 28 x 28 matrix. The architecture of the convolutional neural network consists of multiple layers, such as convolutional layers, max-pooling or average-pooling layers, and fully-connected layers. These different layers convert the raw input pixels into a class score. Overall, our neural network is composed of 12 layers of convolution, ReLU, batch normalization, and global average pooling. The resulting architecture is a fully convolutional neural network with three depthwise-separable residual convolutions, each followed by a batch normalization operation and a ReLU activation function. To generate a prediction, the final layer employs global average pooling and a soft-max activation function. Once the model was trained, we integrated the realized CNN into the NAO robot using the pre-trained model.

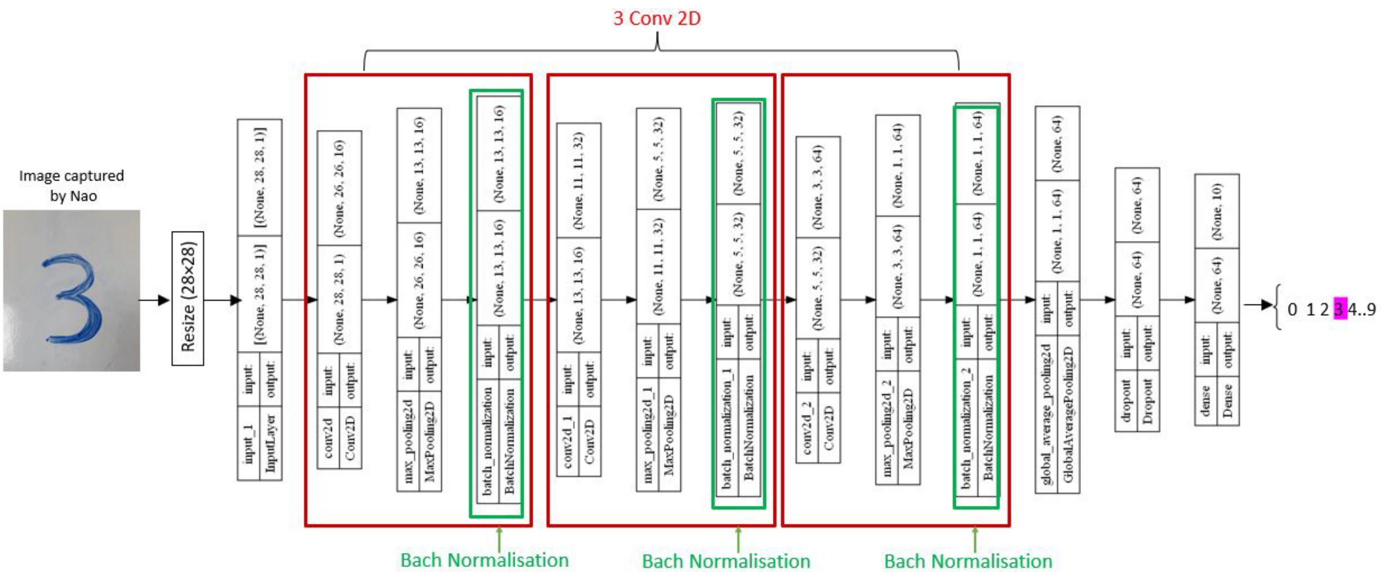



**Figure 5. The model used for HWDR**

## *TESTING THE NETWORK*

The recognition of handwritten digits is a crucial component of the prototype and is the main factor that determines its performance. Therefore, it is imperative to ensure a high level of accuracy of the prototype before deploying it in the real world. The accuracy of the models generated by HWDR was evaluated to ensure that the prototype meets the required standards. As shown in Figure 6, the CNN architecture was trained for 50 epochs, during which the accuracy continuously improved until it reached a remarkable accuracy rate of 99%.

```
42000/42000 [==============================] - 165s 4ms/step - loss: 0.0040 - acc: 0.9985 - val_loss: 0.0522 - val_a
cc: 0.9913
Epoch 45/50
42000/42000 [==============================] - 165s 4ms/step - loss: 0.0053 - acc: 0.9981 - val_loss: 0.0571 - val_a
cc: 0.9899
Epoch 46/50
42000/42000 [==============================] - 165s 4ms/step - loss: 0.0034 - acc: 0.9989 - val_loss: 0.0601 - val_a
cc: 0.9905
Epoch 47/50
42000/42000 [==============================] - 165s 4ms/step - loss: 0.0048 - acc: 0.9984 - val_loss: 0.0566 - val_a
cc: 0.9903
Epoch 48/50
42000/42000 [==============================] - 165s 4ms/step - loss: 0.0052 - acc: 0.9983 - val_loss: 0.0561 - val_a
cc: 0.9908
Epoch 49/50
42000/42000 [==============================] - 165s 4ms/step - loss: 0.0048 - acc: 0.9982 - val_loss: 0.0545 - val_a
cc: 0.9904
Epoch 50/50
42000/42000 [==============================] - 166s 4ms/step - loss: 0.0038 - acc: 0.9988 - val_loss: 0.0530 - val_a
cc: 0.9908
```

**Figure 6. Example of the CNN training**

To evaluate the performance of our developed CNN on each output class, we utilized the confusion matrix, which is a more comprehensive mode of evaluation that provides more information. The confusion matrix highlights correct and incorrect predictions, which are then divided by class, allowing for a detailed analysis of the model's performance. Figure 7 depicts the confusion matrix, which provides a clear view of our prediction results. It is evident from the matrix that the model performed well on most of the classes, with a few misclassifications between similar digits such as 4s and

9s. Overall, the confusion matrix demonstrates that our CNN model is capable of accurate handwritten digit recognition with a high level of precision.

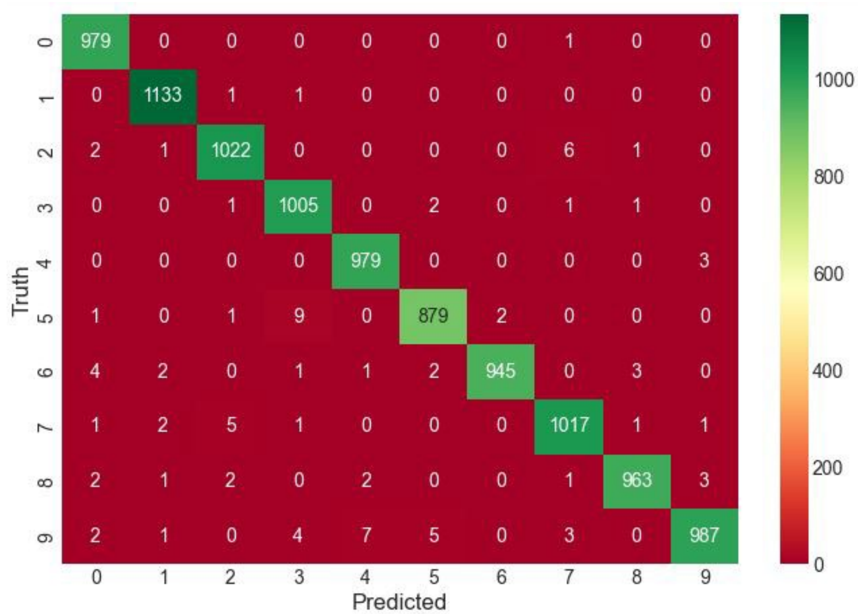



**Figure 7. Confusion matrix**

All of the diagonal elements are correct predictions, for example, we predicted the digit zero, 979 times correctly. The red cell, value shows the wrong predictions. A confusion matrix is a table often used to evaluate the result of a predicate classifier on a test data set with known true values. Furthermore, some performance metrics, such as precision, recall, and F-factor, are used to assess the classifier's performance (Gouraguine et al., 2022). These three performance measures are calculated from the confusion matrix. Table 3 demonstrates our neural network model's classification ratio. Although our CNN model's accuracy could achieve 99%, we were still able to evaluate our results using the classification ratio and enhance our dataset. We have obtained a precision, recall, and F-factor of 99%.

**Table 3. Classification report**

| CLASS | **Precision** | **Recall** | **F1-Score** | **Support** |
|---|---|---|---|---|
| 0 | 0.99 | 1.00 | 0.99 | 980 |
| 1 | 0.99 | 1.00 | 1.00 | 1135 |
| 2 | 0.99 | 0.99 | 0.99 | 1032 |
| 3 | 0.98 | 1.00 | 0.99 | 1010 |
| 4 | 0.99 | 1.00 | 0.99 | 982 |
| 5 | 0.99 | 0.99 | 0.99 | 892 |
| 6 | 1.00 | 0.99 | 0.99 | 958 |
| 7 | 0.99 | 0.99 | 0.99 | 1028 |
| 8 | 0.99 | 0.99 | 0.99 | 974 |
| 9 | 0.99 | 0.98 | 0.99 | 1009 |
| **Accuracy** | | | 0.99 | 10000 |
| **Macro Avg** | 0.99 | 0.99 | 0.99 | 10000 |
| **Weighted Avg** | 0.99 | 0.99 | 0.99 | 10000 |

We compared the accuracy and execution time using experimental graphs for better comprehension and model consistency verification. Figure 8 shows the accuracy obtained during training and validation. As seen, the model achieved a validation loss of 0.04 and an accuracy of 0.99 at epoch 11, indicating that it had learned the patterns in the training set well. This result demonstrates the effectiveness of our model in accurately recognizing handwritten digits.

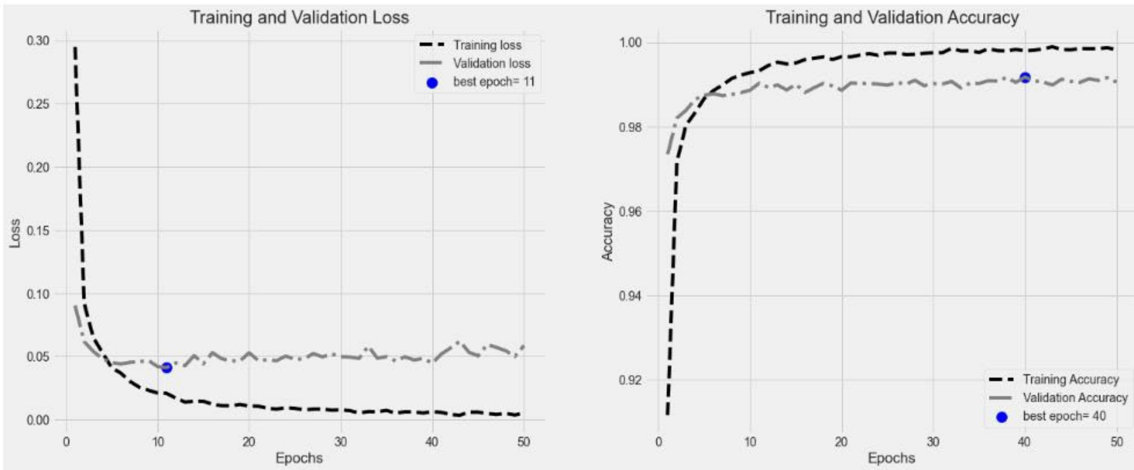



**Figure 8. Training and testing function**

## *SIMULATING THE HANDWRITING DIGIT RECOGNITION*

The HWDR model, which was developed using a CNN, has been effectively implemented in real-time image recognition with the NAO robot. As shown in Figure 9, a suitable environment has been created to ensure smooth interaction between the user and the NAO robot. To ensure the accuracy of the recognition, the handwritten digit must be placed directly in front of the robot's face. The NAO robot has two HD cameras that enable it to recognize faces and objects, which is an essential feature for our recognition model. Additionally, its advanced multimedia system includes four microphones and two speakers that allow for voice synthesis and the announcement of prediction results. The combination of these features and the HWDR model's effectiveness allows for a perfect interaction between the user and the NAO robot, making it an ideal tool for tasks such as educational or assistive technology.

**Figure 9. Environment setup**

The screenshots of the successfully classified digits by our CNN and the corresponding output of the simulated robot are presented in this part. Figure 10 represents the result of the prediction of digit two captured by the robot in real time. The robot pronounces the digit detected and recognized by saying "It is the digit two." We have reinforced our investigations with a student-robot interaction so that students with learning disabilities can overcome and confront their special needs through our new primitive, implemented and deployed on the assistant-tutor robot NAO. Figure 11, shows the recognition of the different digits and the output of each of them by the NAO robot. The robot can even detect if no digits are available on the board, Figure 12 shows the result.

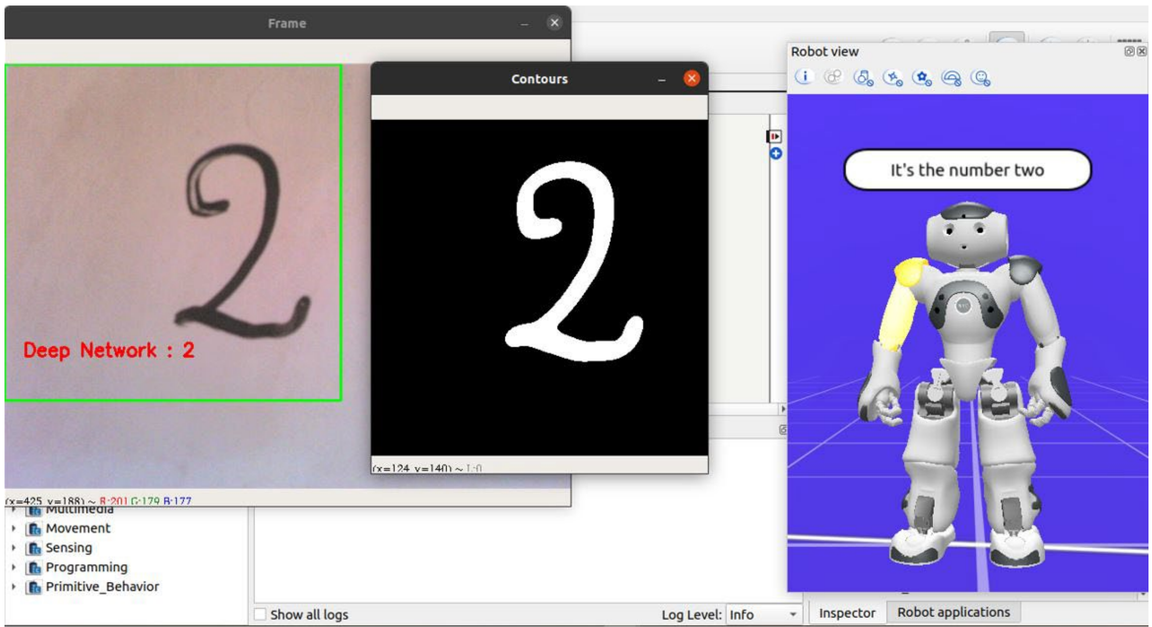
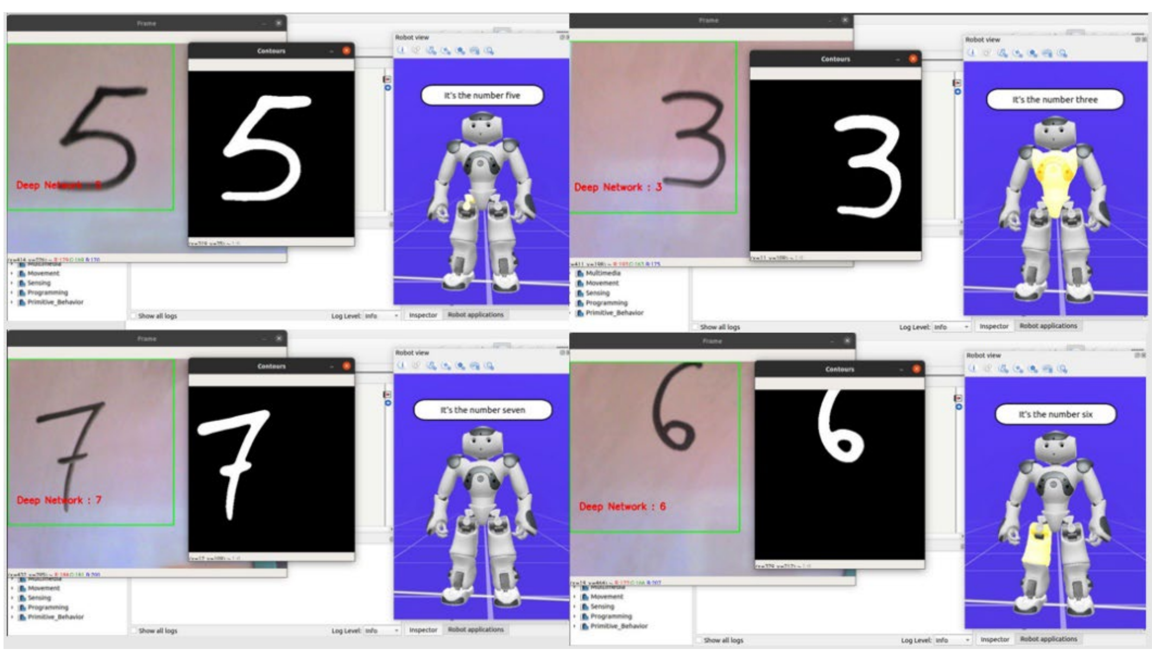



**Figure 10. Example of recognition of the number two**



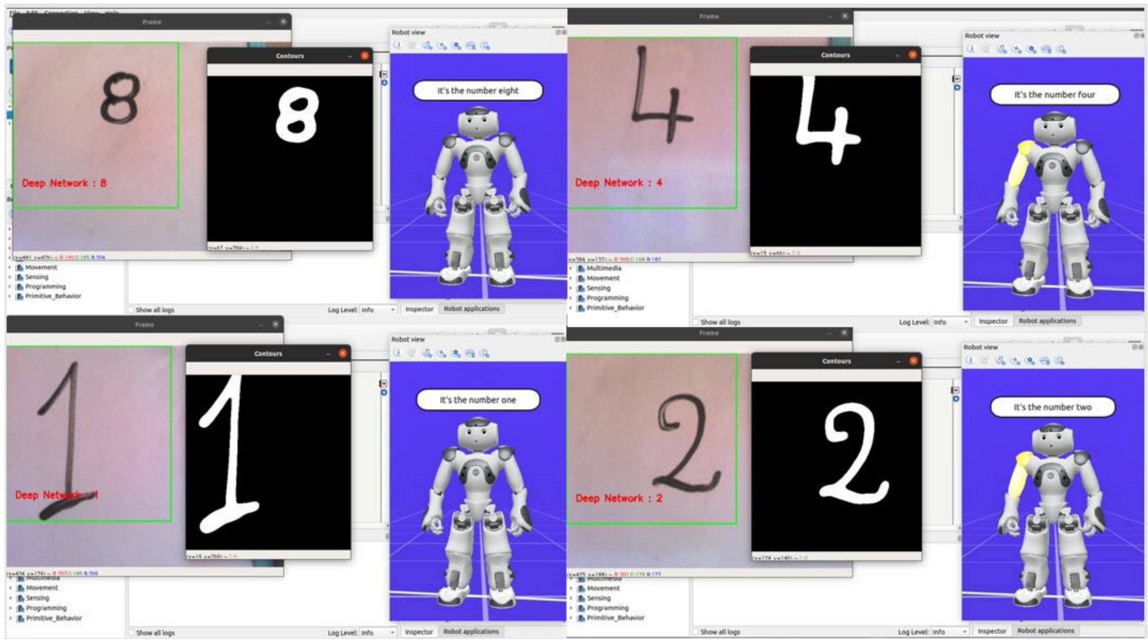



**Figure 11. Examples of execution and output for all digits**

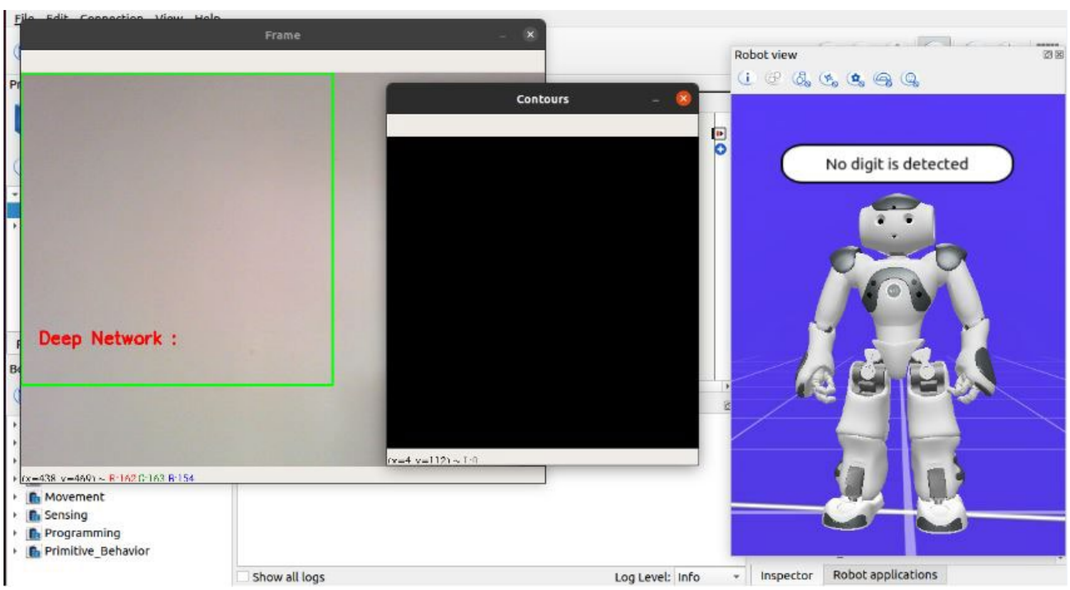



**Figure 12. Example of no digit detected**

## *The Effectiveness of Visual Learning Using the NAO Robot*

The evaluation stage is grounded in a comprehensive methodology that focuses on enhancing the visual learning experiences of children with specific needs. We aimed to facilitate the understanding and interpretation of mathematical concepts, particularly handwritten digits, through engaging interactions involving the NAO robot. These interactions were thoughtfully designed, incorporating group sessions for collaborative learning and individual interactions for personalized attention. During group interactions, students collaborated with the robot, engaging in discussions and exploring mathematical content. In individual sessions, students had the opportunity to present their handwritten digits to the robot, fostering a unique interactive experience.

The learning strategies for which activities have been produced are the visual learning strategy based on the use of the NAO robot (Teaching based on Assistive Technologies: TAT), the visual learning strategy using traditional visual aids such as maps, images, and videos (Teaching based on Traditional aids: TT), and the visual learning strategy combining the NAO robot and traditional visual aids (Both Together - BT). A questionnaire focusing on these three strategies was presented on a sheet of paper adapted for children was distributed to the participants after the teaching activity had been completed. Participants answered the questionnaire according to their level of satisfaction. This questionnaire uses a Likert scale where the values (1, 2, 3, 4, 5) are represented by stars. Here are some examples of the questions we asked:

- How many stars would you give NAO's intelligence?
- How effectively do you think the NAO robot recognized your handwritten digits?
- How engaging did you find the learning activities with the NAO robot?
- How well did the NAO robot assist you in explaining and justifying your numerical answers?
- How likely are you to recommend using the NAO robot for learning to your peers?
- How well did the NAO robot's explanations and demonstrations help you in understanding?
- Do you like using things like pictures, maps, or videos to learn?
- Do you like learning with NAO better than using pictures, maps, or videos to learn?

Figure 13 shows the preferences for teaching and learning methods between male and female students, regardless of their specific needs. The results of the questionnaire revealed that the first strategy, involving the use of the NAO robot, was the best-received, and most effective strategy adopted by the participants. This result confirms the positive impact of integrating the NAO robot into the visual learning process. Participants expressed a higher level of engagement, enjoyment, and understanding when interacting with the NAO robot. They appreciated its ability to recognize handwritten numbers and found this method of teaching approved by the robot intriguing. These results validate the effectiveness of modern technologies and assistive tools like robots in enhancing children's visual learning, improving their cognitive abilities, and stimulating their critical thinking skills.

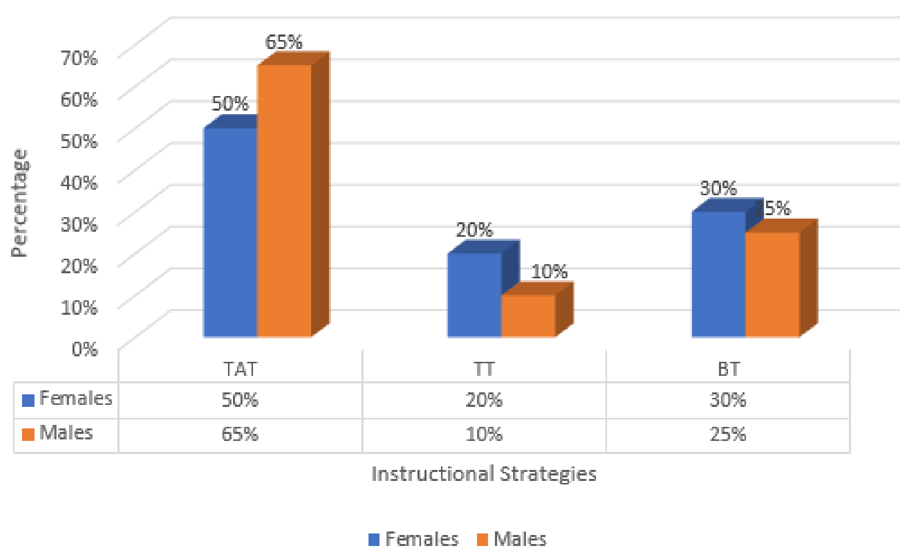



**Figure 13. Distribution according to the preference of instructional strategies most suitable for male and female students**

## DISCUSSION

In this study, we propose a distributed system to enhance the visual learning of children and young students with specific needs by providing the NAO robot with a cognitive intelligence capability (Qidwai et al., 2020). This cognitive intelligence system is designed to enable real-time recognition of handwritten digits, a crucial aspect of visual learning (Dixit et al., 2020). Our approach leverages a new model of Handwritten Digit Recognition (HWDR), which forms the basis of a new knowledge primitive integrated into the NAO robot's functionalities (Filippini et al., 2021). By seamlessly integrating these elements, our distributed system holds the potential to revolutionize the way visual concepts are comprehended and learned by students during their interactions with the NAO robot.

The prototype is distributed across four different computing nodes, with the master server acting as the middleware between the NAO robot and the processing nodes. The deep learning model runs on the master server and performs different pre-processing operations to make real-time predictions. We evaluated the effectiveness of our approach by focusing on a specific learning strategy. Our prototype, which uses a deep learning model based on the HWDR, achieved a high accuracy rate of 99% in recognizing students' handwritten digits during real-time interactions with the NAO robot.

To assess the practical implications and pedagogical effectiveness of our prototype integrated into the NAO robot, we conducted an experiment based on supervised learning sessions by students with special needs (Lesort et al., 2020). Guided by the specific learning strategy of our concept, the students engaged in collaborative group activities and personalized one-to-one interactions with the robot (Shamsuddin et al., 2017). During these sessions, observations were made to capture the students' commitment, reactions, and interactions with the robot. Students demonstrated a high level of enthusiasm for mathematical concepts, actively engaged in discussions with the robot, and asked relevant questions. One-to-one interactions revealed great interest from students in presenting their handwritten digits to the robot. Some students used gestures and corporal expression to communicate, enriching their learning experience.

These observations generated valuable qualitative data, highlighting the students' active engagement and positive interaction with the robot. In addition, the questionnaire administered to the child participants after the activity provided crucial insights into their reactions and perspectives. These findings greatly contributed to a comprehensive evaluation of our robot NAO enhanced visual learning approach, including its effectiveness and how students experienced the interaction with the robot in terms of comprehension and learning. The analysis of the results indicated that the most appropriate distribution according to the preference of teaching strategies for students was Teaching Based on Assistive Technologies (TAT), reinforcing the importance of assistive technologies in the context of inclusive education (Yousif & Yousif, 2020).

## CONCLUSION

Our research aims to incorporate several programming techniques that involve object detection, data acquisition, and real-time data processing with modules of the NAO robot to create a harmonious system for recognizing handwritten digits. Our approach used for the creation and implementation of handwritten digit recognition primitive in our NAO robot is an interesting one, based on the use of distributed systems that divide the work using various components distributed over several nodes, coordinating their efforts to perform tasks more efficiently than a single device.

The successful integration of CNN into our NAO robot significantly improved classification accuracy, showcasing its human-level performance in handwritten digit recognition. To rigorously evaluate the effectiveness of our developed approach, we established an enriched learning environment and conducted engaging activities using the NAO robot as a powerful tool to facilitate visual comprehension and enhance the engagement of learners with specific needs. The results consistently demonstrate the efficacy of our visual learning strategy, highlighting NAO's potential as an impactful tool for tailored education.

This research represents a significant contribution to the advancement of knowledge in the field of improving visual learning in children with special needs, opening up new perspectives for the evolution of educational practices. These findings highlight the promising potential of the NAO robot as a powerful tool for enhancing visual learning experiences, surpassing the effectiveness of traditional teaching methods. The study has certain limitations, particularly because of the small number of participants. Indeed, the recruitment of learners with specific needs is a complex and challenging process, which led to a relatively small sample of twelve children aged between 5 and 7 years old. Despite this, our research succeeded in demonstrating the effectiveness of using our NAO robot to support visual learning in these children in a convincing way.

The promising results of this study confirm the relevance of future research aimed at further exploring the potential of social robots, to enhance student learning and development inclusively and equitably. In our future work, we aim to use several robots that will be executing tasks simultaneously where the considerable increase and growing complexity of the related tasks would be impossible to manage for a single machine. This makes our distributed computing approach very usable and also offers additional advantages over traditional computing environments. The findings of our research can serve as an application for the use of different instructional methods that can be implemented to support the learning needs of all learners.

## REFERENCES


Acharya, B., & Sahu, K. (2020). Software development life cycle models: A review paper. *International Journal of Advanced Research in Engineering and Technology*, *11*(12), 169–176.

Alam, M., Vidyaratne, L., Wash, T., & Iftekharuddin, K. M. (2016, March). Deep SRN for robust object recognition: A case study with NAO humanoid robot. *SoutheastCon 2016*, *Norfolk, VA, USA*, 1–7. https://doi.org/10.1109/SECON.2016.7506751

Albani, D., Youssef, A., Suriani, V., Nardi, D., & Bloisi, D. D. (2017). A deep learning approach for object recognition with NAO soccer robots. In S. Behnke, R. Sheh, S. Sarıel, & D. Lee (Eds.), *RoboCup 2016: Robot World Cup XX* (pp. 392–403). Springer. https://doi.org/10.1007/978-3-319-68792-6_33

Al-Khresheh, M. H., Khaerurrozikin, A., & Zaid, A. H. (2020). The efficiency of using pictures in teaching speaking skills of non-native Arabic beginner students. *Universal Journal of Educational Research*, *8*(3), 872–878. https://doi.org/10.13189/ujer.2020.080318

Aly, A. A., & Dugan, J. B. (2017, September). Experiential robot learning with deep neural networks. *Proceedings of the 7th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics, Lisbon, Portugal*, 356–361. https://doi.org/10.1109/DEVLRN.2017.8329830

Atanga, C., Jones, B. A., Krueger, L. E., & Lu, S. (2020). Teachers of students with learning disabilities: Assistive technology knowledge, perceptions, interests, and barriers. *Journal of Special Education Technology*, *35*(4), 236–248. https://doi.org/10.1177/0162643419864858

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, *3*(21). https://doi.org/10.1126/scirobotics.aat5954

Bouck, E. C., & Long, H. (2021). Assistive technology for students with disabilities: An updated snapshot. *Journal of Special Education Technology*, *36*(4), 249–257. https://doi.org/10.1177/0162643420914624

Cohen, G., Afshar, S., Tapson, J., & van Schaik, A. (2017, May). EMNIST: Extending MNIST to handwritten letters. *Proceedings of the International Joint Conference on Neural Networks*, *Anchorage, AK, USA,* 2921–2926. https://doi.org/10.1109/IJCNN.2017.7966217

Cruz, N., Lobos-Tsunekawa, K., & Ruiz-del-Solar, J. (2018). Using convolutional neural networks in robots with limited computational resources: Detecting NAO robots while playing soccer. In H. Akiyama, O. Obst, C. Sammut, & F. Tonidandel (Eds.), *RoboCup 2017: Robot World Cup XXI* (pp. 19–30). Springer. https://doi.org/10.1007/978-3-030-00308-1_2

Desideri, L., Negrini, M., Malavasi, M., Tanzini, D., Rouame, A., Cutrone, M. C., Bonifacci, P., & Hoogerwerf, E.-J. (2018). Using a humanoid robot as a complement to interventions for children with autism spectrum disorder: A pilot study. *Advances in Neurodevelopment Disorders*, *2*, 273–285. https://doi.org/10.1007/s41252-018-0066-4

Dhillon, A., & Verma, G. K. (2020). Convolutional neural network: A review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, *9*(2), 85–112. https://doi.org/10.1007/s13748-019-00203-0

Dixit, R., Kushwah, R., & Pashine, S. (2020). Handwritten digit recognition using machine and deep learning algorithms. *International Journal of Computer Applications*, *176*(42), 27–33. https://doi.org/10.5120/ijca2020920550

Edwards, C., Edwards, A., Spence, P. R., & Lin, X. (2018). I, teacher: Using artificial intelligence (AI) and social robots in communication and instruction. *Communication Education*, *67*(4), 473–480. https://doi.org/10.1080/03634523.2018.1502459

Filippini, C., Perpetuini, D., Cardone, D., & Merla, A. (2021). Improving human-robot interaction by enhancing NAO robot awareness of human facial expression. *Sensors*, *21*(19), 6438. https://doi.org/10.3390/s21196438

Gouraguine, S., Riad, M., Qbadou, M., & Mansouri, K. (2022). Dysgraphia detection based on convolutional neural networks and child-robot interaction. *International Journal of Electrical and Computer Engineering*, *13*(3), 2999–3009. https://doi.org/10.11591/ijece.v13i3.pp2999-3009

Heinmäe, E., Leoste, J., Kori, K., & Mettis, K. (2022). Enhancing teacher-students' digital competence with educational robots. In M. Merdan, W. Lepuschitz, G. Koppensteiner, R. Balogh, & D. Obdržálek (Eds.), *Robotics in education* (pp. 155–165). Springer. https://doi.org/10.1007/978-3-030-82544-7_15

Henschel, A., Hortensius, R., & Cross, E. S. (2020). Social cognition in the age of human–robot interaction. *Trends in Neurosciences*, *43*(6), 373–384. https://doi.org/10.1016/j.tins.2020.03.013

Jung, S. E., & Won, E. S. (2018). Systematic review of research trends in robotics education for young children. *Sustainability*, *10*(4), 905. https://doi.org/10.3390/su10040905

Karayaneva, Y., & Hintea, D. (2018a). Object recognition algorithms implemented on NAO robot for children's visual learning enhancement. *Proceedings of the 2018 2nd International Conference on Mechatronics Systems and Control Engineering* (pp. 86–92). ACM. https://doi.org/10.1145/3185066.3185083

Karayaneva, Y., & Hintea, D. (2018b). Object recognition in Python and MNIST dataset modification and recognition with five machine learning classifiers. *Journal of Image and Graphics*, *6*(1), 10–20. https://doi.org/10.18178/joig.6.1.10-20

Lesort, T., Lomonaco, V., Stoian, A., Maltoni, D., Filliat, D., & Díaz-Rodríguez, N. (2020). Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges. *Information Fusion, 58*, 52-68. https://doi.org/10.1016/j.inffus.2019.12.004

Munna, A. S., & Kalam, M. A. (2021). Teaching and learning process to enhance teaching effectiveness: A literature review. *International Journal of Humanities and Innovation*, *4*(1), 1–4. https://doi.org/10.33750/ijhi.v4i1.102

Neumann, M. M. (2020). Social robots and young children's early language and literacy learning. *Early Childhood Education Journal*, *48*, 157–170. https://doi.org/10.1007/s10643-019-00997-7

Pot, E., Monceaux, J., Gelin, R., & Maisonnier, B. (2009, September). Choregraphe: A graphical tool for humanoid robot programming. *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication, Toyama, Japan,* 46–51. https://doi.org/10.1109/ROMAN.2009.5326209

Qidwai, U., Kashem, S. B. A., & Conor, O. (2020). Humanoid robot as a teacher's assistant: Helping children with autism to learn social and academic skills. *Journal of Intelligent and Robotic Systems, 98*, 759–770. https://doi.org/10.1007/s10846-019-01075-1

Schnekenburger, F., Scharffenberg, M., Wülker, M., Hochberg, U., & Dorer, K. (2017, November). Detection and localization of features on a soccer field with feedforward Fully Convolutional Neural Networks (FCNN) for the adult-size humanoid robot sweaty. *Proceedings of the 17th IEEE-RAS International Conference on Humanoid Robots, Birmingham, UK.* http://lofarolabs.com/events/robocup/ws17/papers/Humanoids_RoboCup_Workshop_2017_pape_4.pdf

Shamsuddin, S., Azhar, U. N. A. M., Zulkifli, W. Z., & Kamat, S. R. (2017, September). Kansei emotions on human-robot interaction with humanoid robot NAO at kindergarten setting. *Proceedings of the International Conference on Design and Concurrent Engineering, Osaka, Japan.* https://www.researchgate.net/publication/319875778_Kansei_Emotions_on_Human-robot_Interaction_with_Humanoid_Robot_NAO_at_Kindergarten_Setting

Tamakloe, D. (2020). Enhancing learning and development of young children with disabilities with assistive technology. In D. Chambers (Ed.), *Assistive technology to support inclusive education* (pp. 141–161). Emerald Publishing. https://doi.org/10.1108/S1479-363620200000014012

vom Brocke, J., Hevner, A., & Maedche, A. (2020). Introduction to design science research. In J. vom Brocke, A. Hevner, & A. Maedche (Eds.), *Design science research. Cases* (pp 1–13). Springer. https://doi.org/10.1007/978-3-030-46781-4

Woo, H., LeTendre, G. K., Pham-Shouse, T., & Xiong, Y. (2021). The use of social robots in classrooms: A review of field-based studies. *Educational Research Review*, *33*, 100388. https://doi.org/10.1016/j.edurev.2021.100388

Yousif, M. J., & Yousif, J. H. (2020). Humanoid robot as assistant tutor for autistic children. *International Journal of Computation and Applied Sciences*, *8*(2), 8–13. https://www.researchgate.net/publication/342052032_Humanoid_Robot_as_Assistant_Tutor_for_Autistic_Children

Yuan, F., Klavon, E., Liu, Z., Lopez, R. P., & Zhao, X. (2021). A systematic review of robotic rehabilitation for cognitive training. *Frontiers in Robotics and AI*, *8*, 605715. https://doi.org/10.3389/frobt.2021.605715

Zeaiter, S. (2020). Using humanoid robots as tools for the promotion of STEM-education in schools (RoboPraX). In S. Zeaiter & P. Franke, P. (Eds.), *Pioniergeist, ausdauer, leidenschaft. Festschrift zu ehren von Prof. Dr. Jürgen Handke* (pp.279-320). Philipps-Universität Marburg. https://www.researchgate.net/publication/342521772_Using_humanoid_robots_as_tools_for_the_promotion_of_STEM-education_in_schools_RoboPraX

Zhexenova, Z., Amirova, A., Abdikarimova, M., Kudaibergenov, K., Baimakhan, N., Tleubayev, B., Asselborn, T., Johal, W., Dillenbourg, P., CohenMiller, A., & Sandygulova, A. (2020). A comparison of social robot to tablet and teacher in a new script learning context. *Frontiers in Robotics and AI*, *7*, 99. https://doi.org/10.3389/frobt.2020.00099

## AUTHORS

**Soukaina Gouraguine** is a Ph.D. student and part of the team in the Distributed Computing Systems in the research laboratory of Signals, Distributed Systems, and Artificial Intelligence at the ENSET Institute of Mohammedia. Her doctoral work explores the creation and integration of intelligent robotic agents for assistance in innovative educational systems, especially for users (students, teachers, ...) with special needs. She can be contacted at email: soukainagoura@mail.com

**Mohammed Qbadou** obtained a master's degree in Mechanical Engineering from ENSET of Mohammedia in 1992, and a DEA in Energetics and Physics in 1993. He obtained his first Ph.D. in Robotics in the modeling and control of flexible manipulator robots at the Mohammed V University of Rabat in 1998, an HDR in Computer Science in 2017, and his second Ph.D. in Computer Science in 2021 at the Hassan II University of Casablanca. Since 1998, he has been a research professor in Computer Science at ENSET of Mohammedia. His research focuses on the Semantic Web, Big Data Analytics, Artificial Intelligence, Inclusive Smart Education Systems, and Assistive Robotics. He has accumulated 30 years of experience in teaching mechanical engineering, robotics, and computer science. In scientific research, he has produced over 100 indexed publications. He can be contacted at email: qbmedn7@gmail.com

**Mohamed Rafik** works as a researcher and teacher in Electrical Engineering and Embedded and Mobile Intelligent Systems at the ENSET Institute of Hassan II University of Casablanca. His research interests include Energy Efficiency and Renewable Energies, Information Systems, Artificial Intelligence, Inclusive Intelligent Educational Systems, and Assistive Robotics. He graduated from ENSET Mohammedia in 1990 in Electronics, earned a Master's degree in Distributed Information Systems from ENSET Mohammedia, and obtained a Ph.D. in Electrical Engineering, Energy Efficiency, and Renewable Energies from Hassan II University of Casablanca in 2021. He has accumulated 32 years of experience in teaching electrical engineering and computer science. He can be contacted via email: rafik@enset-media.ac.ma

**Mustapha Riad** is a Ph.D. student and part of the team of Distributed Computing Systems in the research laboratory of Signals, Distributed Systems, and Artificial Intelligence at the ENSET Institute of Mohammedia. He received his Master's degree in Information Systems Engineering (ISI) from the Cadi Ayyad University Marrakech in 2019. His doctoral work explores the contribution to the development of intelligent education systems and adaptive learning through the personalization and adaptation of educational content for the learner. He can be contacted at email: My.mustapha.riad@gmail.com

**Khalifa Mansouri** works as a researcher and computer science teacher at the ENSET Institute of the University Hassan II of Casablanca. His research interests include Real-Time Systems, Information Systems, e-Learning Systems, and Industrial Systems (Modeling, Optimization, and Numerical Computing). He has degrees from ENSET of Mohammedia in 1991, CEA in 1992, and a Ph.D. in calculation and optimization of structures at Mohammed V University in Rabat in 1994, HDR in 2010, and a Ph.D. in computer science from the University of Hassan II in Casablanca in 2016. He can be contacted at email: khmansouri@hotmail.com