IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Verification of Student Uniform by Convolutional Neural Network through Images

Long Huynh Hoang
High School Student, FPT High School, Can Tho City, Vietnam
Linh Nguyen Hoang Anh
Falcuty of Interdisciplinary Sciences, Ho Chi Minh City, Vietnam
linhnhahcmus.edu.vn

Abstract

Uniform verification holds critical significance in various sectors like law enforcement, healthcare, and hospitality, ensuring security and professionalism. Manual inspection methods often exhibit inefficiencies and inaccuracies, necessitating automated solutions. This project proposes a novel Convolutional Neural Networks (CNNs) approach for real-time uniform detection and authentication. By harnessing deep learning algorithms, our system achieves exceptional accuracy, notably 91%, enhancing efficiency and mitigating manual inspection burdens. Extensive experimentation validates the efficacy of our CNN-based solution, showcasing its potential for diverse deployments. Beyond enhancing uniform verification reliability, our system lays groundwork for advancing AI-driven object recognition, security surveillance, and automated inspection systems. This research contributes to evolving AI applications, fortifying security and trust within uniformed environments, and paving the way for automated authentication and verification systems' future developments.

Keywords: Uniform verification, Convolutional Neural Networks, Image classification, Clothes Detection.

1 Introduction

School uniforms have been present for a long time in countries worldwide as a distinctive cultural feature within educational environments. In Vietnam, school uniforms have also been in existence for a consider- able period, particularly in major cities, gradually becoming widespread across the country. The act of students wearing uniforms not only signifies a cultural beauty within educational institutions but also encap- sulates various meanings. Uniforms foster a sense of camaraderie among students, demonstrating a spirit of unity within a collective. Therefore, adhering to the uniform policy has become a mandatory regulation in Vietnamese schools.

Artificial Intelligence (AI) is a forefront techno- logical domain, crucial in aiding humans to address numerous issues across various sectors such as manu- facturing, healthcare, education, and more. It stands out for its ability to learn autonomously and enhance accuracy through self-gathering data without requir- ing explicit programming. Relying solely on object features, AI can provide precise identification and clas- sification results through real-life image capture.

Currently, the inspection of student uniforms is manually conducted using the naked eye. This process consumes a lot of time, effort, and productivity is not high. Applying technology, especially artificial intelli- gence, to automate the inspection of student uniforms is an optimal solution to address the current issue. Therefore, we have developed an algorithm to recog- nize and detect whether students are wearing their uniforms correctly or not. Real-world evaluations have shown that the accuracy of the algorithm reaches 91% and it can be further developed for practical applica- tions.

The proposed algorithm introduces several novel as- pects:

1. Firstly, we employ a Convolutional Neural Net- work (CNN) model for object recognition. The advantage lies in the compact size of the trained model, facilitating easy integration on mobile de- vices such as smartphones, tablets, etc.

2. Secondly, we utilize a database collected in Viet- nam from real-life images of students wearing uniforms to enhance the algorithm's accuracy and suitability in our country.

3. Thirdly, the algorithm can perform real-time recognition and classification, providing instan- taneous results.

The following outline will be used for the paper: In

Section II, we will discuss related studies. Sections III and IV will respectively present and evaluate the effec- tiveness of the proposed algorithm. Finally, Section V will conclude and propose directions for future devel- opment.

2 Related paper

Currently, AI has been well-developed to optimize pro- cesses for real-world applications, including its applica- tion in education. The most prevalent research nowa- days focuses on preparing the education system to adapt to the advancement of AI. Both teachers and students require a level of understanding of AI, as well as computer literacy at a high level. On the other hand, there is still a scarcity of research on leveraging AI for optimization within the education system, par- ticularly in schools, especially in Vietnam. However, AI has been developed to support the educational pro- cess in Europe. According to Holmes and Tuomi, there are three deployment forms of AI in the education sys- tem in Europe: AI targeting students, AI targeting teachers, and AI targeting schools (Holmes and Tuomi, 2022) [1] To effectively manage student information and ensure compliance with regulations within schools, several management support systems have been devel- oped:

- RFID-based Student Information Management System (Saleh Alghamdi, 2019) [2] In this study, the author constructed a model comprising hard- ware, including RFID transceivers integrated into student cards. Each student's ID card con- tains personal information, and when students swipe their ID cards on the signal receiver, the system records their school attendance informa- tion.

- Facial Recognition-based Student Attendance System (Alvin Sarraga Alon, Cherry D. Ca- suat)[3].This method utilizes the YOLOv3 neu- ral network model embedded in a computer sys- tem with integrated cameras to identify students through facial recognition, automating the atten- dance process.

- School Uniform Detection and Classification Sys- tem [4]: In this research, the author employed the YOLOv5 neural network model. However, the processing speed and accuracy of the model were not high. Moreover, due to the large size of the model, it is challenging to embed it on mobile devices such as smartphones and tablets.

In the aforementioned methods, a common issue is the relatively low processing speed. Although the RFID- based student information management system is quite accurate, it relies on expensive hardware systems and cannot detect or store information about student uni- forms at the time of inspection. Additionally, many studies have applied image processing technology using neural network models such as YOLOv5, AlexNet, Caf- feNet, etc., to perform recognition without the need for specialized hardware. However, these methods demand computer systems with high configurations and can- not be integrated into low-configured mobile devices. Moreover, the database of student uniforms in Viet- nam is not diverse and is limited in quantity, resulting in the relatively low accuracy of the recognition model. To address the aforementioned challenges, we pro- pose utilizing a Convolutional Neural Network (CNN) model. Data will be collected from schools across Viet- nam in sufficient quantities. Subsequently, data aug- mentation algorithms will be employed to introduce noise factors into the recognition target, enhancing ro- bustness. Finally, the data will be labeled, resized, and meticulously classified before training. By lever- aging the CNN model, the algorithm's processing speed will be optimized, and the size of the algorithm model will be minimized, suitable for hardware configurations

ranging from low to high.

3 Process

3.1 Overall Process

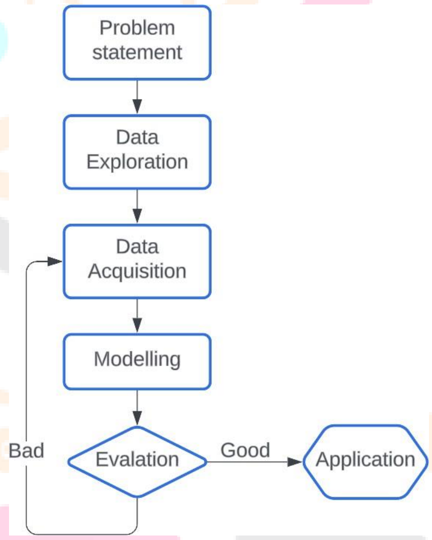

Figure 1: Overall process To

be more specific:

- Problem statement: Observe, study, and identify issues within the local school environment.

- Data exploration: Explore existing data and de- termine the scope of the obtained data.

- Data acquisition: Collect data by capturing im- ages under various environmental conditions and from different individuals. Then preprocess the data and label it as follows:

- Uniform: Brown long pants + Orange collar T-shirt / Knee-length skirt + Orange collar T-shirt.
- Non-uniform: Other types of clothing

- Modeling: Apply and build models for training based on a 7:3 ratio for training and testing.

- Evaluation: Assess the model using evaluation metrics such as accuracy, loss function, and confusion matrix.

  - If the result is ”Good”: Proceed with real-world testing and integration onto mobile devices.
  - If the result is ”Bad”: Enhance data collec- tion and algorithm improvement.

3.2 Data set

During the data collection process, we opted for a manual approach by directly capturing images from mobile and digital devices worn by individuals under diverse lighting conditions, angles, and colors for la- beling as ”Uniform”. While this method consumed time initially, the uniqueness of each school’s uniform and the variations in color, length, and specific regula- tions posed a challenge. Hence, actively collecting data through manual image capture allowed us to maintain the highest control over the quality of the data used.

- Illustration of labeled ”Uniform” data - Students wearing school uniforms (orange T-shirt), and uniform pants (brown trousers):

Figure 2: Examples of students wearing uniform

- Illustration of labeled ”Non-uniform” data - Students wearing various materials and types of clothing, such as T-shirts, shirts, sweaters, shorts, jeans, leggings, etc., in different colors. They may wear the correct top but incorrect bot- toms, or vice versa.

Figure 3: Examples of students not wearing uniform

The dataset was collected through a methodical pro- cess that involved obtaining consent from each student before taking photographs, both in school uniform and regular attire. We aimed to attract a diverse range of student subjects, considering factors such as different locations, lighting conditions, camera angles, etc. Ad- ditionally, images from the school’s social media plat- forms were incorporated to enrich the dataset further. After collection, we amassed a total of 1466 images, comprising 939 images of students in uniforms and 527 images of students not in uniforms. These images were processed to ensure uniformity in size and for- mat. Subsequently, the dataset was split into training and testing sets, with a 70% and 30% ratio for train- ing and testing, respectively. In the experimental set, images were labeled as ”uniform” or ”non-uniform” for evaluation purposes. To prepare the data for machine learning, all images were resized to a standard size of 150x150 pixels and converted into arrays. An addi- tional dimension was added to the array to meet the requirements of deep learning models. We also normalized the pixel values to ensure consistency across the dataset, facilitating effective model training and evaluation.

3.3 Theory Basis

Deep learning: a subset of machine learning that uses multi-layered neural networks, called deep neural networks, to simulate the complex decision-making power of the human brain through a combination of data in- puts, weights, and bias [6]. These elements work to- gether to accurately recognize, classify, and describe objects within the data.. Some form of deep learning powers most of the artificial intelligence (AI) in our lives today. Convolutional Neural Network (CNN): a category of machine learning model, namely a type of deep learning algorithm well suited to analyzing vi- sual data. CNNs – sometimes referred to as convnets – use principles from linear algebra, particularly con- volution operations, to extract features and identify patterns within images. Although CNNs are predomi- nantly used to process images, they can also be adapted to work with audio and other signal data. CNN archi- tecture is inspired by the connectivity patterns of the human brain - in particular, the visual cortex, which plays an essential role in perceiving and processing vi- sual stimuli [7]. The artificial neurons in a CNN are arranged to efficiently interpret visual information, enabling these models to process entire images. Because CNNs are so effective at identifying objects, they are frequently used for computer vision tasks such as im- age recognition and object detection, with common use cases including self-driving cars, facial recognition and medical image analysis.

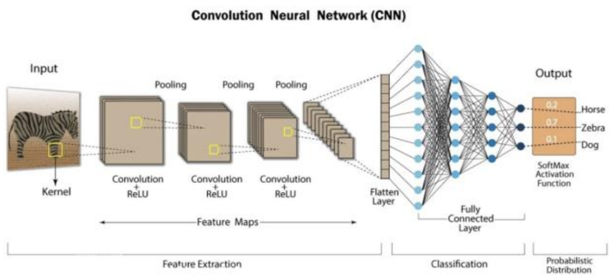

Figure 4: Process of Convolution Neural Network (CNN) [8]

The use of Convolutional Neural Networks (CNNs) for a project such as image recognition and classifica- tion offers several advantages [5]:

- Capability to learn high-level features from data: The Convolutional layers of CNNs automatically extract complex features from images such as edges, corners, colors, or specific object charac- teristics without human intervention. This en- ables the model to learn good feature representa- tions from the data, while reducing dependence on manual feature extraction techniques.
- Position invariance: CNNs can recognize objects in images without knowing the exact position of those objects in the image. Convolutional lay- ers reduce the number of required parameters by weight sharing, and Pooling layers help reduce the size of feature representations without losing important information. Therefore, CNNs can ef- ficiently process images of different positions and sizes.

3.4 Modeling

During the research phase to identify suitable ap- proaches for the project, we prioritized both process- ing speed and overall accuracy of school uniform de- tection. With this aim in mind, we opted for a CNN model with predefined basic metrics. Consequently, we constructed a neural network model using the Keras library with a TensorFlow backend to perform image recognition tasks using Convolutional Neural Networks (CNNs) for classifying images into two groups: contain- ing the target object to be classified or not.

Below is a detailed description of each step in the process:

1. Selecting the CNN model: using the Keras li- brary, we construct a neural network model em- ploying a CNN architecture. The neural network comprises Convolutional layers (Conv2D) com- bined with MaxPooling2D layers to extract im- age features. Subsequently, the Flatten layer is utilized to convert from a 2D tensor to a 1D vec- tor, which is then connected to Dense layers (fully connected layers) for classification.

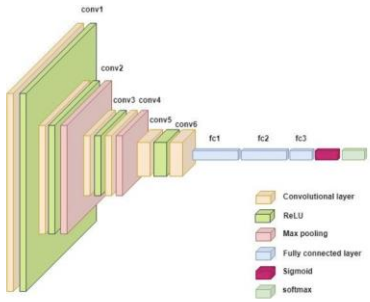

Figure 5: A 16-layer convolutional neural network (CNN) [9]

2. Integra involves defining directories for storing training and testing data and mounting them into the code file. We use the ImageDataGenerator to generate appropriate image data for model train- ing and testing. The data is loaded using the flow from directory method, specifying the input image size, batch size, and other parameters.

```
# Define parameters
batch_size = 32
epochs = 20
IMG_HEIGHT = 150
IMG_WIDTH = 150
```

3. Establishing training parameters: ¡image¿ Defin- ing the necessary parameters for the training pro- cess such as batch size, number of epochs, and input image size
4. Evaluating the loss function after each training iteration. Using the compile method to compile the model with the selected loss function and op- timization algorithm. Conducting the training process with the fit method and assessing the model's performance on the test data after each epoch using the evaluate method.

Figure 6: Training and Validation Loss Graph

In the first two epochs, the loss function steadily decreases to 0, and after completing more than 2 epochs, it remains stable at 0 until the end of the training process. This indicates the stability of the model and signs of not being overfit during training.

4 Result

1. Accuracy: In the training and testing process with a 7:3 ratio and approximately 2000 images in total, we obtained an accuracy score of 0.88 for the first training session. However, recog- nizing the potential for further improvement, we performed data augmentation and collected ad- ditional data, resulting in a total of over 9000 im- ages. From this, we achieved a better accuracy score of 0.92 (92%).

2. Image of the results obtained when performing a trial run on the model:

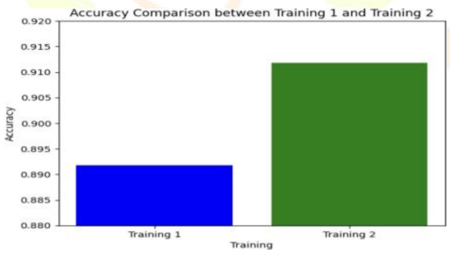

Figure 7: Accuracy comparison between Training 1 and 2

Here, we can observe that in both images, the two students are standing in a relaxed posture. This indicates the basic behavior of students when be- ing photographed, either in front of a camera or in daily life, which fundamentally does not af- fect the process of detecting clothing and mak- ing correct or incorrect determinations regarding uniforms. We believe this is a significant step for- ward in further development for the future as we continue to refine the model for direct camera- based inspections.

Figure 8: Result of testing images

3. Model size:

- Parameter Count: 413696 (parameter) x 4 (bytes/float32) = 1654784 bytes = 1.58 MB
- Input Data Size: 300000000 (pixels) x 4 (bytes/float32) = 1200000000 bytes = 1.12 GB

=¿ Overall Project Size: summing the model size and the input data size, the project's total size is 1.58 MB + 1.12 GB = 1.12 GB.

4. Computational Efficiency: Employing the latest CNN architecture contributes to enhanced com- putational efficiency, with the model achieving a processing speed of 8.989012e-10 seconds, or ap- proximately 0.000000889 seconds.

5 Discussion

In the overall evaluation, the model proposed in the re- lated work segmented the clothing into separate parts, detecting the correctness of each part individually, and then merged the results to calculate the overall accu- racy. The accuracy of the upper part was found to be 0.8, while for the lower part, it was 0.82, resulting in an overall accuracy of approximately 0.81. On the other hand, our model (AL) combines both pants and shirts for evaluation. This decision was made because a school uniform is considered valid only when both pants and shirts are valid. Therefore, we opted to de- tect both pants and shirts as a unified entity. As a result, the accuracy of our model surpassed 0.9, which is a significant improvement. This achievement opens up new possibilities for the field of artificial intelligence, paving the way for future interdisciplinary projects.

Regarding size, the model proposed in the re- lated work appears to be bulky and requires high- performance laptops to operate, with a size of 48.3MB. In contrast, our research focuses on developing a lightweight model with high suitability that can run on various device platforms, particularly smartphones.

Our model occupies only 1.58MB, which is significantly smaller, being 31 times smaller than the previous one. This reduction in size makes our model more versatile and accessible across different hardware configurations, aligning with the trend towards mobile and edge com- puting.

In terms of runtime, perhaps the most signifi- cant and prominent difference between our approach and theirs lies in the time it takes to process the data. Since we use a Convolutional Neural Network (CNN) for training, while they utilize YOLOv5 for the same purpose, the difference in speed is quite evident. They take approximately 5-6 seconds or 5000-6000 mil- liseconds to provide prediction results after capturing an image, whereas our model can produce results in just 8.989012e-10 seconds or 0.000000889 milliseconds. This stark contrast in processing time represents a sig- nificant advancement and could be considered a sub- stantial step forward for future developments in this area.

Furthermore, due to limited conditions at the school, their dataset consists of only 1500 images, not to mention that all the students in the images are fe- male, resulting in low diversity and accuracy of the model. In contrast, given favorable conditions, we de- voted all our dedication and efforts, spending over four months to collect and meticulously curate the data for our dataset. Moreover, we applied various augmenta- tion techniques to enhance the diversity of the data. As a result, our dataset comprises over 9400 images divided into two classes - uniform and non-uniform at- tire - captured from both male and female students, from various angles, and under different lighting con- ditions. All these efforts were aimed at creating the best possible model and striving for the model's future suitability and development.

6 Conclusion

In conclusion, the research has proposed a convolu- tional neural network (CNN) model to detect and rec- ognize students wearing school uniforms correctly or incorrectly. The proposed algorithm achieves an accu- racy rate of over 91%, with a high processing speed of 8.989012e-10 (s) and a relatively small size of 1.58MB, while not requiring high-performance hardware con- figuration. These results enable the algorithm to be easily integrated into mobile devices such as smart- phones, tablets, etc. However, algorithm development still faces challenges due to the diversity of school uni- forms in each school and country. This necessitates us to continuously update new data and optimize the training process. School uniforms represent a cultural beauty in every school and country. Through this re- search, we aim to contribute to the development of ed- ucation, promote the application of technology in ed- ucation, especially AI technology. Additionally, we as- pire to diversify the global database of school uniforms, thereby enhancing the accuracy of AI technology in the future.

7 References

[1] Holmes, W ., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542-570. https://doi.org/10.1111/ejed.12533

*[2]*Saleh Alghamdi (2019). MONITORING STU- DENT ATTENDANCE USING A SMART SYSTEM AT TAIF UNIVERSITY. International Journal of Computer Science & Information Technology (IJCSIT) Vol 11, No 1, February 2019

*[3]*Alvin Sarraga Alon , Cherry D. Casuat. A YOLOv3 Inference Approach for Student Attendance Face Recognition System. International Journal of Emerging Trends in Engineering Research.

*[4]*J. Park, H. Lee, M. Choi and E. Shin, "Uniform Detection Using YOLOv5", GitHub, Mar. 24, 2024. (accessed Apr. 01, 2024)

*[5]*J. Teuwen and N. Moriakov, "Convolutiona=l neural networks," Handbook of Medical Image Com- puting and Computer Assisted Intervention, pp. 481–501, 2020.

*[6]*IBM, "What is Deep Learning?," www.ibm.com, 2023. https://www.ibm.com/topics/deep-learning

*[7]*R. Awati, "What is convolutional neural net- work? - Definition from WhatIs.com," SearchEnter- priseAI, Sep. 2022.

*[8]*Haque, Kh.N. (2023) What is Convolutional Neural Network — CNN (Deep Learning).

*[9]*Sree Hari Rao . G. Madhu. (2017) Malaria Eradication Strategies Program (MESP):Challenges and Information Technology(IT)PerspectivesV.