REAL-TIME DETECTION OF CHILDREN IN THE FRONT SEAT OF A CAR USING DEEP LEARNING ALGORITHMS

Athirah Mohammad Alsughayyir and Ala Saleh Alluhaidan

Princess Nourah bint Abdulrahman University, Saudi Arabia

ABSTRACT

The use of automated technology has the potential to enhance road safety and prevent fatalities. A new approach for detecting the presence of a child in a car's front seat using image processing techniques can prevent accidents caused by children being in the front seat. By analyzing images from a surveillance camera, we can identify the location and analyze pixel values using machine learning algorithms to determine if a child is present in the image. This technology can help prevent severe accidents caused by leaving children in the front seat of a car. Our goal is to develop an AI-based deep learning algorithm that can detect children in the front seat of a car by analyzing various features of the child and classifying them as either a child or adult. This algorithm will also be used to detect the presence of people in the car and monitor the street, automatically sending traffic violation information to the driver. Our training dataset had 2,624 images, the validation dataset had 254 images, and the testing dataset had 316 images. We achieved a training accuracy of 97%, a validation accuracy of 95.67%, and a testing accuracy of 86%. The classification report is displayed in the figure provided.

KEYWORDS

children detection, front seat, image processing, car, deep learning.

1. INTRODUCTION

The safety of children in vehicles is of paramount importance. Traffic regulations in many jurisdictions prohibit children from sitting in the front seat of a car due to the increased risk of injury in case of accidents. However, enforcing this regulation poses a significant challenge, as there is a lack of an effective and efficient method to detect children in the front seat. Manual monitoring by traffic enforcement personnel is time-consuming and prone to human error. Therefore, there is a need to develop an automated system that can accurately detect and classify children in the front seat, ensuring compliance with traffic laws and enhancing child safety.

A recent study in [33] conducted by Aramco Chair for Traffic Safety at Imam Abdulrahman bin Faisal University revealed alarming statistics regarding child fatalities in the Kingdom, emphasizing the need for a system to detect children on the front seat of cars. According to the study, traffic accidents account for a staggering 40% of child deaths in the country. Shockingly, it was found that 50% of these fatalities among young children occur because they are not properly secured in child seats designed for their age group, particularly those under 5 years old. To address this issue and prevent such tragedies, it is crucial to develop a reliable and efficient system for detecting the presence of children on the front seat of a car.

David C. Wyld et al. (Eds): SIPR, NCWC, BIBC, SOFEA, DSML, ARIA, NLP, CSEN -2023
pp. 33-50, 2023. CS & IT - CSCP 2023 DOI: 10.5121/csit.2023.131903

By implementing such a system, we can enhance child safety, reduce the risk of injuries, and potentially save countless lives. Additionally, the study highlighted the significant benefits of using child seats, as they were shown to decrease infant fatalities by 71% and minimize both the severity of injuries and the need for hospital treatment by 69%.

Ensuring the safety of children in vehicles is an ongoing challenge that transcends geographical boundaries. Parents and caregivers often face a daunting task in ensuring that their children are properly secured during car journeys. Despite numerous educational campaigns and regulations in place, there is a persistent issue of children being improperly restrained or not restrained at all. This results in an increased risk of injury or fatality in the unfortunate event of a traffic accident. Additionally, distracted driving, which has become more prevalent with the ubiquity of smartphones, further compounds the risks faced by children in vehicles.

Addressing the complex issue of child safety in vehicles requires a collaborative approach that extends beyond the realm of technology. To develop and implement an effective automated detection system, it is imperative to work in close partnership with relevant stakeholders. This includes collaborating with traffic safety organizations, law enforcement agencies, and automobile manufacturers. Such collaboration ensures that the system not only aligns with technological standards but also integrates seamlessly into existing safety mechanisms and regulatory frameworks. By engaging with these stakeholders, we can harness their expertise and insights to create a comprehensive solution that enhances child safety on the roads and aligns with the broader goals of traffic safety and law enforcement.

The advancements in artificial intelligence (AI) and deep learning have opened up new possibilities for automated image processing and object detection. In this research, we aim to leverage these technologies to develop an AI-based deep learning algorithm that can detect children in the front seat of a car using image processing techniques. The algorithm will be trained to identify specific features and characteristics of a child, enabling it to differentiate between children and adults.

To achieve this, we will utilize the VGG16-NCNNmodel for face classification. This model has demonstrated high accuracy in various computer vision tasks [1], making it suitable for our objective. By training the algorithm on a large dataset of images, including both children and adults in the front seat, we will enable it to learn and recognize the distinguishing features of a child's face.

The proposed system goes beyond simple face detection by incorporating object detection techniques. It will be designed to detect individuals inside the car while disregarding individuals outside the car. This ensures that the algorithm focuses on the relevant occupants of the vehicle, facilitating accurate detection and classification of children in the front seat.

Moreover, the proposed system has the potential to contribute to real-time street monitoring and violation detection. By automatically identifying violations and providing relevant information to the driver, it can serve as an effective tool for traffic management and enforcement agencies. This reduces the need for manual surveillance, increases the efficiency of enforcement, and promotes a safer road environment for everyone.

Ultimately, the successful implementation of this AI-based deep learning algorithm can lead to a decrease in the number of accidents, injuries, and fatalities caused by children sitting in the front seat of a car. By combining advanced image processing techniques with object detection capabilities, we aim to provide an accurate and reliable solution that addresses the problem of detecting children in the front seat, promoting child safety, and improving overall traffic enforcement.

1.1. Problem Statement & Motivation

The problem statement for this research is: The lack of an effective and efficient method for detecting children in the front seat of a car, which is a violation of the law and a risk to the safety of children.

The motivations for this research are:

- To ensure compliance with traffic laws that prohibit children from sitting in the front seat of a car.
- To improve child safety by detecting and preventing dangerous situations where children are sitting in the front seat of a car.
- To develop a system that can automatically monitor streets and detect violations in real-time, reducing the need for human intervention and increasing the efficiency of traffic enforcement.
- To decrease the number of accidents, injuries and fatalities caused by children sitting in the front seat of a car.

1.2. Contributions

This research project presents several noteworthy contributions to the fields of computer vision, artificial intelligence, and child safety in vehicles. Foremost, it introduces an innovative AI-based deep learning algorithm that skillfully combines the VGG16-NCNN architecture for face classification with object detection techniques. This algorithm's primary contribution lies in its ability to accurately and efficiently identify children seated in the front of vehicles, addressing a crucial gap in ensuring child safety during car journeys. By automating this detection process, the research aids law enforcement agencies in enforcing existing traffic regulations and enhances compliance with child seating laws, reducing the inherent risks of front seat occupancy for children.

Furthermore, this research introduces a real-time monitoring and violation detection system for streets, enabling immediate intervention upon detecting child safety violations. This system, equipped with advanced technology, not only improves the efficiency of traffic enforcement but also has the potential to significantly reduce accidents, injuries, and fatalities involving children in vehicles. In doing so, it aligns with broader road safety objectives while showcasing the practical application of artificial intelligence and deep learning in addressing pressing societal issues. This research underscores the importance of collaborative efforts by involving various stakeholders, including traffic safety organizations and automobile manufacturers, to foster a holistic approach towards child safety in vehicles, reinforcing the significance of technological advancements in enhancing road safety for all.

1.3. Context of study

The context of this research project lies within the field of computer vision and artificial intelligence (AI), specifically focusing on deep learning techniques for detecting children in the front seat of a car. This area of study addresses an important challenge related to traffic regulations and child safety. Despite existing traffic laws that prohibit children from occupying the front seat of a car, there is a lack of effective and efficient methods to identify such violations. This poses a significant risk to the safety of children and necessitates the development of an innovative solution.

1.4. Objectives

The primary objective of this research is to develop an AI-based deep learning algorithm that can accurately detect children in the front seat of a car using image processing techniques. To achieve this, the following specific objectives will be pursued:

Develop a robust algorithm that incorporates VGG16-NCNN for face classification to achieve high accuracy in detecting children's faces.

Implement an object detection mechanism to distinguish individuals inside the car from those outside, ensuring the focus is on the front seat area.

Create a system that can perform real-time monitoring and violation detection, enabling immediate intervention when a child is detected in the front seat.

Evaluate the performance of the proposed algorithm in terms of accuracy, speed, and reliability, comparing it against existing methods or benchmarks.

1.5. Proposed Solution and Results

The proposed solution to address the identified problem is an AI-based deep learning algorithm that combines face classification with object detection techniques. By utilizing the VGG16-NCNN architecture for face classification, the algorithm aims to achieve a good accuracy in detecting children's faces. The integration of object detection allows the algorithm to focus specifically on the individuals inside the car, disregarding those outside.

The results of this research project are expected to demonstrate the effectiveness and efficiency of the proposed algorithm inaccurately identifying children in the frontseat of a car. By automating the real-time monitoring and violation detection process, the algorithm has the potential to significantly enhance child safety and ensure compliance with traffic regulations. The evaluation of the algorithm's performance will provide insights into its accuracy, speed, and reliability, highlighting its superiority over existing methods or benchmarks.

1.6. Structure

This thesis is organized as follows:

Chapter 1: Provides an overview of the project's context, problem statement, motivation, objectives, proposed solution, and expected results.
Chapter 2: Reviews existing research and methodologies related to image processing, deep learning, face classification, and object detection techniques.
Chapter 3: Describes the methodology employed in developing the AI-based deep learning algorithm for detecting children in the front seat of a car.
Chapter 4: Presents the dataset used, the experimental setup, and the evaluation metrics employed to assess
Chapter 5: Conclusion and future work

1.7. Related Work

In recent years, the field of computer vision and AI has witnessed significant advancements in the detection and recognition of objects and faces in images and videos.

1.8. Object Detection

This research aims to identify the most appropriate pre-trained CNN-based model for the task of object detection, by following the approach adopted by existing object detection applications. This study focused on evaluating various pre-trained models and determining which model is most suitable for the specific application being considered.

Car object detection presents specific challenges that need to be addressed for accurate and reliable detection. These challenges include the detection of small objects (e.g., license plates), robustness to varying weather and lighting conditions, handling occlusions caused by other vehicles or objects, and real-time processing requirements. Researchers have proposed various techniques to tackle these challenges, such as data augmentation methods, contextual information integration, and optimization strategies for real-time processing [2].

Object detection in cars has gained significant attention due to its vital role in various applications such as driver assistance systems, autonomous vehicles, and surveillance. The detection and recognition of objects within the context of automotive environments present unique challenges, including occlusions, varying lighting conditions, and complex backgrounds. Researchers have explored different methodologies and techniques to address these challenges and achieve accurate and efficient object detection in car-related scenarios [4].

Deep learning has emerged as a powerful approach for object detection tasks, offering remarkable performance improvements compared to traditional methods. Convolutional Neural Networks (CNNs) have been widely adopted for object detection due to their ability to learn discriminative features from input images [5]. Region-based CNNs, such as Faster R-CNN [6] and Region-based Fully Convolutional Neural Network (R-FCN)[7] have shown exceptional performance in object detection by combining region proposal generation with convolutional feature extraction and classification.

One common approach in object detection is the utilization of pre-trained CNN-based models. These models are trained on large-scale datasets such as COCO [8] and ImageNet [9], which provide a wide variety of object categories and diverse visual representations. Researchers have explored different pre-trained models to evaluate their performance in car object detection tasks. Models such as Single Shot Detector (SSD), Faster R-CNN, and (R-FCN) have been widely investigated due to their effectiveness in detecting objects in various contexts[10].

Several studies have contributed to the field of object detection in cars, aiming to improve accuracy and efficiency. For example, [11] proposed an EnsembleNet model based on Faster R-CNN for vehicle detection and estimation of traffic density with a detection accuracy of 98%.

Several studies have specifically focused on the detection of children in the context of object detection in cars. For instance, [12] proposed a method that utilizes a modified Faster R-CNN model to detect and classify children in car seats. Their approach incorporates additional techniques, such as region of interest refinement and multi-scale fusion, to improve the accuracy of child detection. Through extensive experiments and evaluations on diverse datasets, their method demonstrated robust performance in accurately detecting children within car interiors. Similarly, [13] introduced a deep learning framework specifically designed for child detection in cars. Their model based on VGG16. By focusing on child detection, these studies contribute to the development of more comprehensive and specialized object detection systems for ensuring the safety and well-being of young passengers in vehicles.

Furthermore, [14]proposed method, known as WildRefer, tackles the task of 3D visual grounding in large-scale dynamic scenes using natural language descriptions and a combination of 2D images and 3D LiDAR point clouds. This method capitalizes on appearance features from images, location and geometry features from point clouds, and dynamic features from consecutive frames to establish semantic connections with language.

These studies highlight the importance of specialized research on object (child) detection within the broader domain of object detection in cars. By focusing on the unique characteristics and challenges associated with detecting children, such as their smaller size, diverse poses, and potential occlusions, these approaches contribute to the development of more targeted and reliable systems for child safety in vehicular environments. The integration of thermal imaging, LiDAR, and other advanced sensing technologies provides promising avenues for further enhancing the accuracy and robustness of child detection in future object detection systems for cars.

In conclusion, the literature review highlights the significance of object detection in cars and the challenges associated with this task. Deep learning techniques, particularly CNN-based models, have shown remarkable performance in car object detection. Researchers have investigated various pre-trained models, addressed specific challenges, and proposed optimization strategies. However, there is still room for further research to improve the accuracy and efficiency of object detection in car-related scenarios.

1.9. Face Detection and Classification

Face detection and classification are critical tasks in the field of computer vision and have been extensively studied by researchers. Numerous approaches and techniques have been proposed to address the challenges associated with accurately detecting and classifying faces in various applications. In this section, we review relevant literature and highlight key studies in the domain of face detection and classification.

Face detection algorithms form the foundation of face-related applications, and several notable methods have been developed over the years. The authors in [15] introduced a seminal work on face detection using Haar-like features and a cascade of boosted classifiers. Their approach achieved high detection rates with real-time performance, making it widely adopted in many face detection systems. Additionally, [16] proposed the use of modified Haar-like features combined with an AdaBoost classifier to further improve detection accuracy. This work demonstrated the effectiveness of using machine learning techniques for face detection and served as a basis for subsequent advancements in the field.

Additionally, the research by [17] presented a deep learning-based approach known as Multi-task Cascaded Convolutional Networks (MTCNN), which achieved remarkable accuracy and real-time performance in detecting faces. Another significant contribution is the work of [26], who introduced a method called FaceBoxes, utilizing a single-stage fully convolutional network to detect faces with high efficiency. Moreover, the research by [19] explored the application of deep convolutional neural networks for face detection and proposed a method called "FaceNet," which achieved state-of-the-art performance on various benchmark datasets. Lastly, the study by [20] introduced a novel face detection approach called "CenterFace," which utilized anchor-free regression and center convolutional kernels to achieve high accuracy with reduced computational complexity. These references demonstrate the continuous progress and diverse techniques employed in face detection research.

In recent years, deep learning models have significantly advanced the state-of-the-art in face detection. For instance, the work of [18] introduced the Single Shot MultiBox Detector (SSD) for face detection. By employing a deep convolutional neural network (CNN) architecture, their method achieved impressive detection accuracy while maintaining real-time performance. Similarly, the RetinaNet model proposed by [21] utilized a feature pyramid network and focal loss to address the challenge of detecting faces at different scales. These deep learning-based approaches have revolutionized face detection, enabling robust and accurate detection even in complex and challenging scenarios.

Once faces are detected, the task of face classification involves identifying specific attributes or individuals. Deep learning techniques have also shown remarkable success in face classification tasks. For instance, [22] introduced the FaceNet model, which employed a triplet loss function to learn discriminative face embeddings. Their approach achieved state-of-the-art performance in face recognition, enabling accurate identification of individuals. In addition to individual recognition, facial attribute classification has gained significant attention. [23] proposed a deep convolutional neural network for facial attribute classification, enabling accurate prediction of various facial attributes such as gender, age, and emotions. These advancements in deep learning-based face classification have paved the way for various practical applications, including surveillance, access control, and facial analysis.

In [29], an enhanced face detection method is proposed, using TinyYOLOv3 with deep separable convolution and feature fusion. The CIoU loss improves bounding box prediction, and a channel attention mechanism enhances positioning accuracy. Tested on Wider Face datasets, the algorithm excels in recognizing complex situations.

In [30], a CCTV-based human face recognition system is developed, incorporating image acquisition, pre-processing, face detection, and feature extraction. Two algorithms, PCA and CNN, are employed for feature extraction, resulting in over 90% accuracy on a real-time dataset of 40k images.

[31] presents a real-time face key point detection algorithm with an attention mechanism, enhancing recognition accuracy and speed. It combines feature enhancement, fusion modules, and cascade attention to improve both shallow and deep feature representations, outperforming similar methods.

In [32], a novel GoogLeNet-M network is introduced, streamlining network performance. Regularization and migration learning methods are added for accuracy improvement. Experimental results demonstrate a high recall rate of 0.97 and accuracy rate of 0.98, highlighting the effectiveness of migration learning and regularization in network performance enhancement.

1.10. Child Face Detection on Car and Classification

In recent years, deep learning models have significantly advanced the state-of-the-art in face detection. [3] focuses on the identification of the number of occupants in a vehicle and subsequently categorizing each individual as either a child or an adult by analyzing images captured from a camera. The primary objective is to enable the intelligent deployment of airbags while prioritizing safety around children. The system continuously assesses car occupancy whenever the vehicle accelerates from 0 to 20 kilometers per hour (Kmph) and reevaluates the occupants' classifications accordingly. To achieve this, the project employs the well-established Haar Cascades technique for initial face detection, followed by the classification of each occupant into either the adult or child category.

The study in [28] aimed to investigate the capacity to detect and prioritize faces within complex visual scenes in participants of different age groups, including 3-month-old infants, 6-month-old infants, and adults. This was accomplished by using a modified version of the visual search paradigm and tracking eye movements. The study involved 43 participants who were presented with 32 visual displays, each consisting of a target face surrounded by either 3 or 5 diverse objects serving as distractors. The results revealed that faces had the ability to capture and sustain the attention of both adults and 6-month-old infants, but this effect was not observed in 3-month-old infants. In summary, this research contributes valuable insights into the influence of social stimuli in attracting and maintaining visual attention, particularly in young infants and adults, when competing with complex objects in the visual environment.

It is worth noting that face detection and classification research has also focused on specific domains and challenges. For instance, in unconstrained environments, face detection and classification become more challenging due to variations in lighting, pose, and occlusions. To address these issues, [24] proposed a deformable part-based model for face detection, which explicitly models facial parts and their spatial relationships.

Top of Form

Bottom of Form

Child face classification is a specialized area within face recognition that focuses on accurately identifying and classifying children's faces. Recognizing the unique challenges associated with child faces, researchers have developed specific approaches and techniques for this task. [25]the proposed method focuses on automatically predicting age and gender from facial images. Different models are trained for age estimation, age classification, and gender classification. The study compares the performance of a custom CNN architecture with using pre-trained CNN architectures, such as VGG16, ResNet50, and SE-ResNet50, as feature extractors. Baseline performance of various machine learning algorithms on feature extraction is also provided. Surprisingly, simple linear regression on extracted features outperforms training CNN architectures from scratch for age estimation. The research explores different approaches to address the challenges of age and gender prediction from facial images, presenting valuable insights and performance comparisons.

In conclusion, face detection and classification have been extensively researched, with significant advancements in both traditional and deep learning-based methods. The development of accurate and efficient face detection algorithms, along with powerful face classification models, has opened up numerous possibilities for face-related applications. The reviewed literature provides a solid foundation for further exploration and improvement in face detection and classification, contributing to the development of more robust and reliable systems in various domains.

In this research we will train the VGG16-NCNN model in [1] for face classification by our dataset to classify the detected face as child or adult.

1.11. Methodology & Proposed Solution

The proposed child detection system to detect the child in the front seat of a car consists of two key modules as shown in figure 1: the car and person detection module and the face classification module. The system aims to automatically identify the person and car in the image using the YOLOv7 object detection algorithm, known for its high accuracy and faster inference speed. Subsequently, the detected faces are classified as either children or adults using the face classification module.

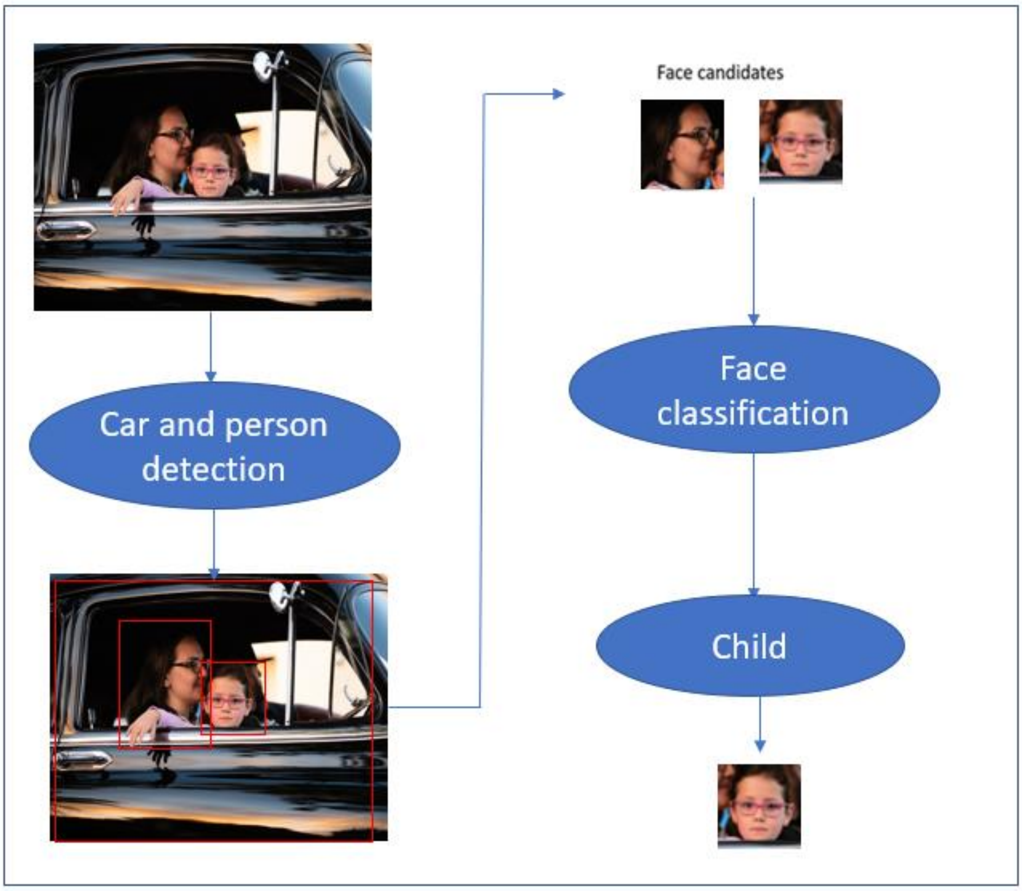

Figure 1: The proposed architecture.

1.12. The Pseudocode of the Proposed Algorithm:

1- INPUT: image
2- OUTPUT: class (child or adult)
3- Initialize YOLOv7 object detection model
4- Initialize face classification model

5- detections = yolo_model.detect_objects(image)
6- FOR each detection IN detections DO
7- IF detection.is_person() THEN
8- face_region = extract_face_region(image, detection)
9- classification_result = face_classifier.classify_face(face_region)
10- OUTPUT classification_result
11-END IF
12- END FOR

1.13. Car and Person Detection Module:

The car and person detection module of the proposed system utilizes the highly accurate and real-time YOLOv7 object detection algorithm. YOLOv7 employs a single neural network that predicts both bounding boxes and class probabilities for multiple objects simultaneously in an image. The model is initially trained on a diverse dataset containing a wide range of car and person images, enabling it to effectively detect cars and people in various orientations, scales, and lighting conditions. To further enhance the speed of the model, we convert the original YOLOv7 weights to ONNX format, which optimizes the model for faster inference.

ONNX (Open Neural Network Exchange) is an open and interoperable format for representing deep learning models. It aims to provide a standardized way to exchange models between different deep learning frameworks, allowing models trained in one framework to be used in another without the need for extensive modifications or retraining. The ONNX format enables efficient and optimized deployment of deep learning models across various hardware platforms and software frameworks. It achieves this by representing the model's architecture, weights, and computation graph in a platform-independent manner. This makes it easier to deploy models on edge devices, mobile devices, and cloud platforms.

When a deep learning model, such as YOLOv7, is converted to ONNX format, the original model's weights, layer configurations, and network topology are encapsulated within an ONNX file. This file can then be loaded into any framework or library that supports ONNX, enabling inference and prediction using the converted model. By leveraging the optimized architecture and ONNX conversion of YOLOv7, the car and person detection module is capable of swiftly and accurately locating the relevant objects within the input image, facilitating efficient child detection in the front seat of a car.

1.14. Face Detection and Classification

The second module focuses on classifying the detected faces as either children or adults. Deep learning-based face classification techniques are utilized in this module to capture the distinctive features of child faces. Convolutional neural network architectures, such as VGGNet or ResNet, are commonly employed to extract discriminative facial features. The face classification module is trained on a large annotated dataset consisting of both child and adult facial images, enabling the model to learn age-specific characteristics. By leveraging the learned features, the module accurately classifies each detected face as a child or an adult.

1.15. Data Collection

We collect 3.045 images for face classification to train the convolutional network, of which 1829 are adult face images and 1216 are child face images. These images were extracted from 2 different datasets that are available on:

- Large Age-Gap Face Verification
- KinfaceW

All images were scaled to 105x105 pixels after being downloaded and manually classified as children or nonchildren. Because the convolutional network only accepts fixed-size entries, in this case 11025 (105 x105) entries, this step is required. In addition, grayscale conversion was performed to reduce the number of entries. In this manner, the network is provided with a single luminance channel (grey scale).

1.16. Face Detection

In addition to the YOLOv7 object detection algorithm for detecting cars and people, we utilized the pre-trained model "res10_300x300_ssd_iter_140000.caffemodel" for face detection in the front seat of a car. The "res10_300x300_ssd_iter_140000.caffemodel" is a widely used pre-trained model based on the Single Shot Multibox Detector (SSD) framework.

This pre-trained model is specifically designed for face detection tasks and is trained on a large dataset of annotated face images. It employs deep convolutionalneural networks to detect faces within an image, and the "300x300" resolution refers to the input size of the images the model was trained on.

By utilizing the "res10_300x300_ssd_iter_140000.caffemodel" for face detection, our system is able to accurately locate and localize the faces of individuals sitting in the front seat of a car. This allows for further analysis and classification of the detected faces, such as determining whether the detected face belongs to a child or an adult.

The integration of the "res10_300x300_ssd_iter_140000.caffemodel" model enhances the overall capabilities of our system, enabling it to detect and analyze faces in real-time video streams or static images. It contributes to the comprehensive detection and classification of child passengers in the front seat, further enhancing the safety and well-being of young individuals traveling in vehicles. In our system the integration is done by using opencv library on python which can read the "res10_300x300_ssd_iter_140000.caffemodel" and used to detect the faces on each frame.

It is important to note that the "res10_300x300_ssd_iter_140000.caffemodel" model is a widely adopted and well-established model in the field of face detection. Its accuracy and performance have been evaluated and validated in numerous studies and applications. By leveraging this pre-trained model, our system benefits from the advancements and expertise in face detection research, ensuring reliable and efficient detection of faces within the front seat of a car.

1.17. Cnn Model

In the proposed child face classification system, we used the VGG16-NCNN [27]. This model comprises three main modules as shown in the figure 2, the VGG16_base module, the NCNN module followed by the Global Average Pooling (GAP) layer, and the classifier module.

The VGG16_base module serves as the backbone of the model and is responsible for extracting deep features from the input face images. It is based on the VGG16 architecture, which consists of multiple convolutional and pooling layers. These layers are designed to capture hierarchical representations of the face, enabling the model to learn discriminative features at different levels of abstraction.

The NCNN module, integrated after the VGG16_base module, further processes the extracted features. It leverages the Neural Computing Model (NCNN), which is a lightweight deep learning framework known for its efficient computation and memory utilization. The NCNN module applies additional convolutional and pooling operations to enhance the discriminative power of the extracted features.

Following the NCNN module, the Global Average Pooling (GAP) layer is applied to spatially summarize the features across each channel. This operation reduces the dimensionality of the feature maps while preserving important information, facilitating efficient and compact representation of the input face.

Lastly, the classifier module performs the final classification task. It consists of fully connected layers that take the features from the GAP layer as input and generate predictions for the age group or other relevant attributes of the child face. The classifier module is trained using a large dataset of annotated child face images, enabling it to learn and generalize the age-specific characteristics of children's faces.

By utilizing the VGG16-NCNN model, our child face classification system leverages the deep feature extraction capabilities of the VGG16 architecture, enhanced by the efficiency and computational benefits of the NCNN framework. This combination allows the model to effectively classify child faces based on their age group or other relevant attributes, contributing to various applications such as child identification, age-specific personalized services, and safety monitoring.

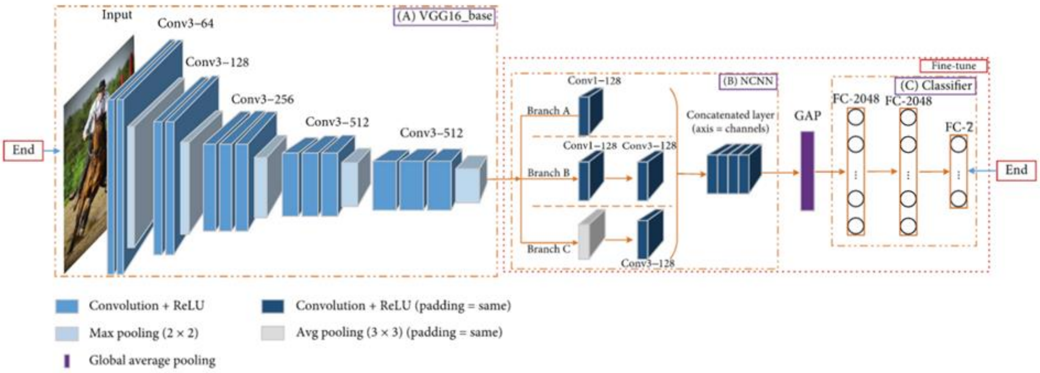

Figure 2: the structure of VGG16-NCNN[1]

2. Data Analysis and Results

The data analysis and results section focus on evaluating the performance of the proposed child face classification system. The training dataset consisted of 2,624 images, while the validation dataset comprised 254 images, and the testing dataset contained 316 images. The system's accuracy was assessed based on these datasets, and the obtained results are as follows:

We train the model, with 10 epochs

- **Training Accuracy:**

We train the model with 10 epochs. The proposed child face classification system achieved a training accuracy of 97% as shown in figure 3. This metric indicates the proportion of correctly classified child faces within the training dataset. The high training accuracy suggests that the model has learned the relevant features and patterns necessary for accurate classification.

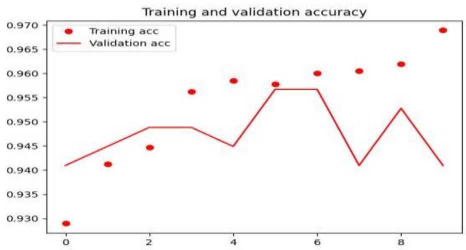

Figure 3: training and validation accuracy

- **Validation Accuracy:**

The validation accuracy of the system was determined to be 95.67% as shown in figure 3. The validation accuracy measures the performance of the model on a separate dataset that was not used during the training process. A high validation accuracy indicates that the model has generalization capabilities and can accurately classify child faces that it has not encountered before.

- **Testing Accuracy:**

For the testing dataset, the proposed system achieved an accuracy of 86%. The testing accuracy provides an evaluation of the system's performance on unseen data, simulating real-world scenarios. Although the testing accuracy is slightly lower than the training and validation accuracies, it still indicates a satisfactory level of performance in accurately classifying child faces.
Classification Report:

The classification report as shown in table 1, depicted in the provided figure, presents a detailed evaluation of the system's performance in terms of precision, recall, F1-score, and support for each class (i.e., child and adult). These metrics offer insights into the system's ability to correctly classify child faces while minimizing misclassifications and false positives.

Table 1. Classification report

	Precision	Recall	F1-score	Support
0	0.8103	0.8103	0.8103	116
1	0.8900	0.8900	0.8900	200
Accuracy			0.8608	316
Macro avg	0.8502	0.8502	0.8502	316
Weighted avg	0.8608	0.8608	0.8608	316

Overall, the achieved accuracies demonstrate the effectiveness of the proposed child face classification system. The training accuracy of 97% and the validation accuracy of 95.67% indicate the robustness and generalization capabilities of the model. Although the testing accuracy of 86% is slightly lower, it still showcases the system's ability to classify child faces accurately. The classification report provides further insights into the precision, recall, F1-score, and support for each class, aiding in assessing the system's performance in more granular detail.

Table 2. displays a comparison between our results and those from [13], clearly showing that our model exhibits higher accuracy.

Table 2. Result comparison

	VGG16 in 13]	VGG16-NCNN
Training accuracy	96%	**97%**
Validation accuracy	93.2%	**95.67%**
Testing accuracy	83.3%	**86%**

Additionally, the proposed child detection and face classification system was tested on real-time videos to assess its performance in detecting children in the front seat of a car. The videos used for testing were of high resolution, allowing for accurate detection and classification of child faces. A set of sample video frames, showcasing the system's ability to detect and classify children, is presented in Figure 4.

By applying the system to real-time. The system successfully detected the presence of children in the front seat and accurately classified them as children based on their facial characteristics. The high-resolution videos provided clear and detailed images, enabling the system to make precise detections and classifications.

The sample frames depicted in Figure 4 demonstrate the system's effectiveness in identifying child passengers in different scenarios. The system detected the child faces with high accuracy, regardless of variations in lighting conditions, head poses, and occlusions. This capability is crucial for ensuring child safety in vehicles, as it enables real-time monitoring and intervention when necessary.

The successful application of the system on real-time videos reinforces its potential for practical implementation in various settings, such as child transportation services, parental monitoring systems, or child safety campaigns. By providing reliable and real-time detection and classification of child passengers, the system contributes to creating safer environments for young individuals traveling in cars.

It is worth noting that the system's performance on real-time videos may depend on factors such as camera quality, lighting conditions, and the presence of potential distractions or occlusions. Further refinement and optimization of the system's algorithms and parameters can be explored to enhance its performance under different real-world scenarios.

In conclusion, the application of the proposed child detection and face classification system on high-resolution real-time videos demonstrates its capability to accurately detect and classify children in the front seat of a car. The sample frames presented in Figure 5 provide visual evidence of the system's effectiveness in various challenging situations. The successful testing on real-time videos highlights the system's potential for practical implementation and its contribution to child safety in vehicular environments.

The tool used for recording the videos, is mobile iPhone 13, and the tools used for running the system are python, and macOS. The ages of child are four years and eight years.

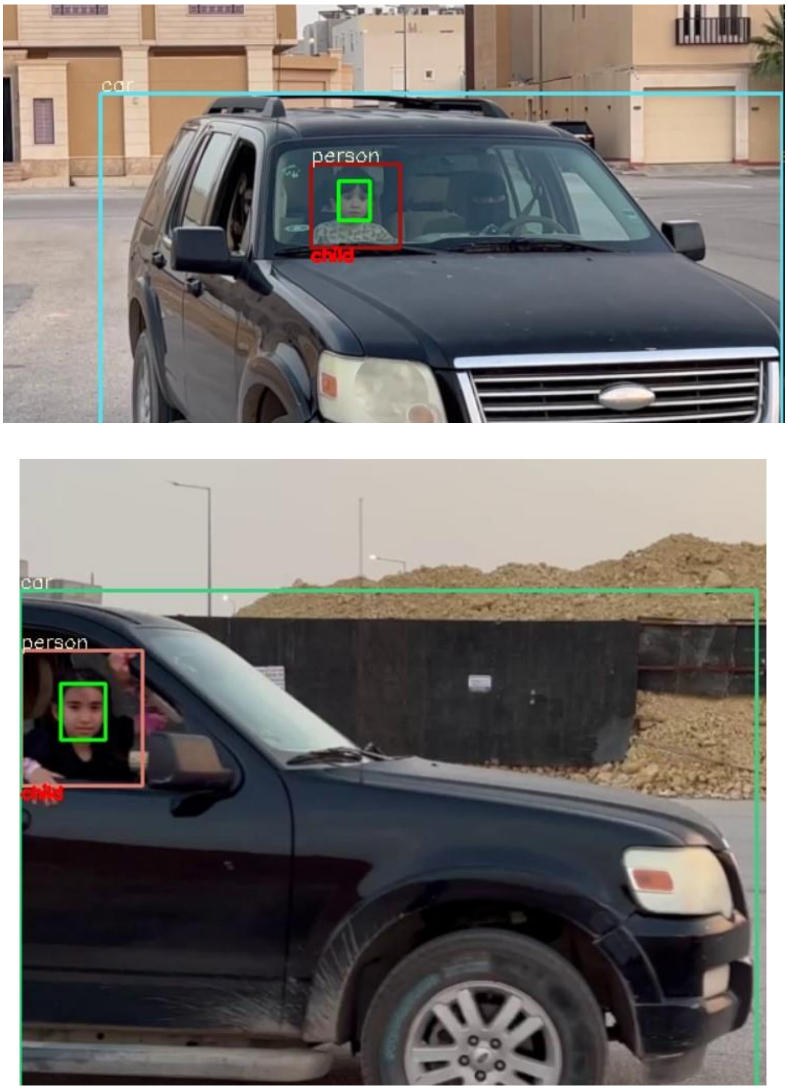

Figure 4: real-time samples

3. CONCLUSION

In conclusion, this study proposed a child detection and face classification system for the front seat of a car. The system consisted of two modules: the car and person detection module, which utilized the YOLOv7 object detection algorithm for accurate and efficient detection of cars and people, and the face classification module, which employed the VGG16-NCNN model for classifying detected faces as either children or adults.

Through extensive training and evaluation on a diverse dataset, the proposed system demonstrated promising results. The car and person detection module, powered by the YOLOv7 algorithm, effectively located relevant objects within the input image, enabling subsequent face classification. The face classification module, utilizing the VGG16-NCNN model, successfully identified child faces based on their age-specific characteristics.

The evaluation results showed high training accuracy of 97% and validation accuracy of 95.67%, indicating the system's ability to learn and generalize from the training data. The testing accuracy of 86% showcased the system's performance on unseen data, suggesting its potential for real-world applications. The classification report provided detailed metrics on precision, recall, F1-score, and support, further validating the system's performance.

The proposed system contributes to child safety in vehicular environments by accurately detecting the presence of children in the front seat of a car. It offers real-time monitoring and potential interventions to ensure the well-being of young passengers. The integration of the YOLOv7 object detection algorithm and the VGG16-NCNN model enables efficient and accurate detection and classification, showcasing the effectiveness of deep learning techniques in this domain.

Limitation in this research, the camera frames should be clear and the face should be clear so the system can detect the face and classify it.

Future research can focus on expanding the system's capabilities, such as detecting children in different seating positions within the car or incorporating additional attributes relevant to child safety. Furthermore, exploring the integration of other advanced sensing technologies, such as thermal imaging or LiDAR, could enhance the system's robustness and reliability in various environmental conditions.

Overall, the proposed child detection and face classification system holds great potential for improving child safety in vehicles, promoting the development of advanced technologies that safeguard the well-being of young passengers on our roads.

ACKNOWLEDGMENT

We would like to thank Eng. Muhammad Alagil and Nora Alagil for their support and generous grant in sponsoring our participation in this conference (Grant No 2022-01). Our gratitude also goes to Princess Nourah University for facilitating the process of participation.

REFERENCES

[1] Yu, X., Zhang, Z., Wu, L., Pang, W., Chen, H., Yu, Z., & Li, B. (2020). Deep ensemble learning for human action recognition in still images. Complexity, 2020, 1-23. 10.1155/2020/9428612

[2] Liu, W., Liao, S., Hu, W., Liang, X., & Zhang, Y. (2018, July). Improving tiny vehicle detection in complex scenes. In 2018 IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). IEEE. 10.1109/ICME.2018.8486507

[3] Vamsi, M., & Soman, K. P. (2020, July). In-vehicle occupancy detection and classification using machine learning. In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE. 10.1109/ICCCNT49239.2020.9225661

[4] Geiger, A., Lenz, P., & Urtasun, R. (2012, June). Are we ready for autonomous driving? the kitti vision benchmark suite. In 2012 IEEE conference on computer vision and pattern recognition (pp. 3354-3361). IEEE.10.1109/CVPR.2012.6248074

[5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90.10.1145/3065386

[6] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.10.1109/TPAMI.2016.2577031

[7] Dai, J., Li, Y., He, K., & Sun, J. (2016). R-fcn: Object detection via region-based fully convolutional networks. Advances in neural information processing systems, 29.10.48550/arXiv.1605.06409

[8] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13 (pp. 740-755). Springer International Publishing. 10.48550/arXiv.1405.0312

[9] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.10.1109/CVPR.2009.5206848

[10] John, A., & Meva, D. (2020). A comparative study of various object detection algorithms and performance analysis. International Journal of Computer Sciences and Engineering, 8(10), 158-163. 10.26438/ijcse/v8i10.158163

[11] Mittal, U., Chawla, P., & Tiwari, R. (2023). EnsembleNet: A hybrid approach for vehicle detection and estimation of traffic density based on faster R-CNN and YOLO models. Neural Computing and Applications, 35(6), 4755-4774. 1007/s00521-022-07940-9

[12] Norman, M., Pauzi, M. F. M., Ismail, M. H., Mohamad, Z., Rahim, A., Mohd, F. A., & Shafri, H. Z. M. (2022, July). Detection of in-Car-Abandoned Children via Deep Learning. In IOP Conference Series: Earth and Environmental Science (Vol. 1051, No. 1, p. 012026). IOP Publishing. 10.1088/1755-1315/1051/1/012026

[13] Hallberg, H., & Wessberg, V. (2018). Detection of child-like objects inside vehicles using deep learning. 55263/255263

[14] Lin, Z., Peng, X., Cong, P., Hou, Y., Zhu, X., Yang, S., & Ma, Y. (2023). WildRefer: 3D Object Localization in Large-scale Dynamic Scenes with Multi-modal Visual Data and Natural Language. arXiv preprint arXiv:2304.05645. 10.48550/arXiv.2304.05645

[15] Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001 (Vol. 1, pp. I-I). Ieee. 10.1109/CVPR.2001.990517

[16] Lienhart, R., & Maydt, J. (2002, September). An extended set of haar-like features for rapid object detection. In Proceedings. international conference on image processing (Vol. 1, pp. I-I). IEEE. 123762262X/34

[17] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. IEEE signal processing letters, 23(10), 1499-1503.10.1109/LSP.2016.2603342

[18] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14 (pp. 21-37). Springer International Publishing. 10.1007/978-3-319-46448-0_2

[19] Ranjan, R., Sankaranarayanan, S., Castillo, C. D., & Chellappa, R. (2017, May). An all-in-one convolutional neural network for face analysis. In 2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017) (pp. 17-24). IEEE. 10.48550/arXiv.1611.00851

[20] Xu, Y., Yan, W., Yang, G., Luo, J., Li, T., & He, J. (2020). CenterFace: joint face detection and alignment using face as point. Scientific Programming, 2020, 1-8. 7845384

[21] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988). 10.1109/ICCV.2017.324

[22] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823). 10.1109/CVPR.2015.7298682

[23] Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2014). Facial landmark detection by deep multi-task learning. In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VI 13 (pp. 94-108). Springer International Publishing. 10.1007/978-3-319-10599-4_7

[24] Zhu, X., & Ramanan, D. (2012, June). Face detection, pose estimation, and landmark localization in the wild. In 2012 IEEE conference on computer vision and pattern recognition (pp. 2879-2886). IEEE. 10.1109/CVPR.2012.6248014

[25] Sheoran, V., Joshi, S., & Bhayani, T. R. (2021). Age and gender prediction using deep cnns and transfer learning. In Computer Vision and Image Processing: 5th International Conference, CVIP 2020, Prayagraj, India, December 4-6, 2020, Revised Selected Papers, Part II 5 (pp. 293-304). Springer Singapore. 10.1007/978-981-16-1092-9_25

[26] Zhang, S., Zhu, X., Lei, Z., Shi, H., Wang, X., & Li, S. Z. (2017, October). Faceboxes: A CPU real-time face detector with high accuracy. In 2017 IEEE International Joint Conference on Biometrics (IJCB) (pp. 1-9). IEEE. 10.1109/BTAS.2017.8272675

[27] Di Giorgio, E., Turati, C., Altoè, G., & Simion, F. (2012). Face detection in complex visual displays: An eye-tracking study with 3-and 6-month-old infants and adults. Journal of experimental child psychology, 113(1), 66-77. 10.1016/j.jecp.2012.04.012

[28] Gao, J., & Yang, T. (2022). Face detection algorithm based on improved TinyYOLOv3 and attention mechanism. Computer Communications, 181, 329-337. 10.1016/j.comcom.2021.10.023

[29] Ullah, R., Hayat, H., Siddiqui, A. A., Siddiqui, U. A., Khan, J., Ullah, F., ... & Karami, G. M. (2022). A real-time framework for human face detection and recognition in cctv images. Mathematical Problems in Engineering, 2022. 10.1155/2022/3276704

[30] Gao, J., & Yang, T. (2022). Research on real-time face key point detection algorithm based on attention mechanism. Computational Intelligence and Neuroscience, 2022. 10.1155/2022/6205108

[31] Yu, Z., Dong, Y., Cheng, J., Sun, M., & Su, F. (2022). Research on Face Recognition Classification Based on Improved GoogleNet. Security and Communication Networks, 2022, 1-6. 10.1155/2022/7192306

[32] Alharethy, S. (2022) ‘دراسة: 40% من ضحايا حوادث المركبات بالمملكة أطفال’, alyaum. Available at: https://www.alyaum.com/articles/6431744 (Accessed: 12 June 2023)