Jurnal Informatika: Jurnal pengembangan IT

Vol. 10, No. 2, 2025

ISSN : 2477-5126, DOI:10.30591/jpit.v10i2.6540

Child Presence Detection for Child Safety with Deep Neural Networks

Sidiq Syamsul Hidayat[1], Dita Aprilia[2], Sindung Hadwi[3], Irfan Mujahidin[4], M. Cahyo Ardi Prabowo[5], Fikri Arif Rakhman[6]

[1,2,3,4,6] *Department of Telecommuncations Engineering, Politeknik Negeri Semarang, Jln. Prof. Sudarto, Tembalang , Kota Semarang, 50275, Indonesia*

[5] *Department of Electrical Engineering, Politeknik Negeri Semarang, Jln. Prof. Sudarto, Tembalang , Kota Semarang, 50275, Indonesia*

Info Artikel

Riwayat Artikel:

Received 2024-01-30
Revised 2025-04-26
Accepted 2025-05-06

Abstract – *Many surveillance systems lack the ability to detect children automatically in environments with minimal supervision. To address this gap, this study proposes a child presence detection system using pretrained deep learning models AgeNet and GenderNet for facial age and gender estimation. These models, based on the Levi and Hassner CNN architecture, are integrated via OpenCV's deep learning module to support real-time processing. The system processes video frames to detect faces, estimate age and gender, and label individuals aged 0–12 as children. Testing was conducted across varying light intensities (110–9440 lux), camera distances (100–500 cm), and heights (70–150 cm). Results show optimal performance at 100 cm distance under outdoor lighting (8410–9440 lux), with detection sensitivity reaching 99.50% for face detection, 96.63% for gender estimation, and 84.72% for child detection. Two-way ANOVA confirmed that light intensity and object distance significantly impact detection performance (p < 0.05), while camera height does not (p = 0.081). This research highlights the practical application of pretrained models in child safety systems. Future work may include IoT integration, real-time alerts, and performance optimization in low-light or dynamic environments.*

Keywords*: Computer Vision; Age Estimation; Gender Estimation; Child-Object Detection; Child Safety*

Corresponding Author:

Sidiq Syamsul Hidayat

Email: sidiqsh@polines.ac.id

Abstrak – *Sebagian besar sistem pengawasan belum mampu mendeteksi keberadaan anak secara otomatis di lingkungan dengan pengawasan terbatas. Untuk menjawab celah ini, penelitian ini mengusulkan sistem deteksi keberadaan anak menggunakan model deep learning pralatih, yaitu AgeNet dan GenderNet, untuk estimasi usia dan jenis kelamin wajah. Model ini, berbasis arsitektur CNN Levi dan Hassner, diintegrasikan melalui modul OpenCV untuk mendukung pemrosesan real-time. Sistem memproses setiap frame video untuk mendeteksi wajah, memperkirakan usia dan gender, lalu memberi label "anak" pada individu berusia 0–12 tahun. Pengujian dilakukan dengan variasi intensitas cahaya (110–9440 lux), jarak kamera (100–500 cm), dan tinggi kamera (70–150 cm). Hasil menunjukkan performa optimal pada jarak 100 cm dengan pencahayaan luar ruangan (8410–9440 lux), dengan sensitivitas deteksi wajah 99,50%, estimasi gender 96,63%, dan deteksi anak 84,72%. Uji statistik ANOVA dua arah menunjukkan bahwa intensitas cahaya dan jarak objek berpengaruh signifikan (p < 0,05), sementara tinggi kamera tidak (p = 0,081). Penelitian ini menyoroti penerapan praktis model pralatih dalam sistem keselamatan anak. Pengembangan selanjutnya dapat mencakup integrasi IoT, notifikasi real-time, serta optimasi performa di kondisi minim cahaya atau lingkungan dinamis.*

Kata Kunci*: Visi Komputer, Estimasi Usia, Estimasi Gender, Deteksi Keberadaan Anak, Keselamatan Anak*

I. INTRODUCTION

Child accidents and injuries can occur unexpectedly in various environments, such as construction zones, industrial facilities, rooftops, or hospital treatment rooms. These risks often result from inadequate adult supervision and underscore the need for intelligent surveillance technologies that can help safeguard children in potentially hazardous areas [1]. Children's health is not only threatened by physical accidents but also by exposure to environments where disease transmission is high. As a result, some local authorities have enacted restrictions, such as limiting access to hospital areas for children under a certain age. However, non-compliance with these regulations remains common, reinforcing the need for automatic child presence detection systems to support policy enforcement [2]. Surveillance cameras have become essential tools for monitoring public and semi-public spaces. With advances in artificial intelligence and computer vision, surveillance systems have evolved from passive monitoring tools into active agents capable of real-time human behaviour recognition and tracking [3].

In the context of improving child safety, several studies have investigated vision-based systems. Naik et al. [2] proposed a proximity hazard detection framework that monitors children near open wells and triggers early warnings to prevent accidents. Dey and Khan [3] designed a deep learning-based solution to differentiate between safe and unsafe objects for children, facilitating better interaction with surrounding environments. Further advancements in facial attribute recognition have enabled accurate age and gender estimation through

CNN-based models. These methods provide a non-intrusive approach to identifying children, which is especially valuable in surveillance scenarios where identity-based recognition is either prohibited or unnecessary.

Islam and Baek [4] developed a deep learning framework capable of estimating real age and gender from unconstrained facial images captured in diverse environments such as retail spaces. Their work utilized large-scale datasets like IMDB-WIKI and Adience, applying CNN models enhanced by image pre-processing techniques to improve classification accuracy. By employing a multitask learning strategy that jointly predicted age and gender, they were able to increase model generalizability in real world settings demonstrating practical potential for integration in surveillance systems. Similarly, Kinthadi et al. [5] focused on predicting human age and gender within smart security systems using a CNN trained on the same benchmark datasets. Their emphasis on real-time deployment and robustness against motion blur and changing backgrounds makes their approach relevant to our objectives. They confirmed that lightweight models could maintain reliable accuracy under constrained computational resources, paving the way for use in embedded surveillance devices.

To address the challenge of real-time age-based child detection, pretrained CNN models such as AgeNet and GenderNet have been widely used. These models classify individuals into predefined age and gender groups using facial features. Deng et al. [6] proposed a multi-feature fusion approach for facial age estimation, combining global and local features to achieve greater accuracy across varied lighting and occlusion levels. Their findings support the use of facial features alone to determine age groups without requiring extensive retraining. Chao et al. [7] implemented pretrained CNNs within a multitask video analytics platform for mobile systems. Their work demonstrated that age and gender models can run simultaneously on limited-resource devices while maintaining acceptable performance—highlighting the feasibility of integrating AgeNet and GenderNet into smart surveillance infrastructure. Similarly, Karatzas and Anagnostopoulos [8] explored the challenges of executing multiple deep neural networks on embedded hardware. Their evaluation showed that pretrained age and gender classifiers can coexist with other real-time models without overwhelming system resources, enabling multitask capabilities in edge devices such as smart CCTV systems. Sun et al. [9] developed a deep conditional distribution learning model for facial age estimation, demonstrating that pretrained CNNs could be used as feature extractors in more sophisticated prediction frameworks. Their method preserved the generalization capability of models like AgeNet while improving adaptability to unseen data distributions—an aspect vital for applications in surveillance where diversity in facial attributes is high.

In this study, we adopt pretrained AgeNet and GenderNet models to estimate the age and gender of individuals captured in video surveillance. These models are deployed using OpenCV's deep learning module to ensure ease of integration and real-time operation. Unlike previous research that involved developing novel network architectures, our approach leverages publicly available models to reduce complexity and improve accessibility. Novelty of this study lies in its focus on implementing, real-time child presence detection system based on age estimation. We evaluate the system's performance across different lighting conditions, camera distances, and mounting heights. Furthermore, statistical analyses using two-way ANOVA are conducted to assess the significance of external variables on detection outcomes. Whereas earlier studies were limited to controlled environments or adult-focused surveillance, this research presents a deployable solution for detecting children in diverse settings, including public spaces and in-vehicle systems. By integrating proven CNN models into a tailored video surveillance pipeline, this work contributes to the ongoing effort to enhance child safety through smart technologies.

II. METHOD

A. System Design

The detection system utilizes a Canon EOS M50 camera to capture video input in MP4 format with a resolution of 1920×1080 pixels at 24 frames per second. These videos are transferred to a Linux-based machine running Ubuntu 16.04 and processed using Python 3.7 and OpenCV's deep learning module (cv2.dnn) [10]. The system performs face detection on each video frame. Once a face is detected, the image region is cropped and resized to 227×227 pixels before being passed to the AgeNet and GenderNet models for age and gender estimation. The estimated outputs are overlaid on the video using bounding boxes and labels, as illustrated in the process flow diagram in Figure 1. Starting with video input, frames are extracted, followed by face detection. Detected faces are preprocessed and passed into AgeNet and GenderNet. The resulting age and gender classifications determine whether the object is a child. If so, the label is rendered on-screen and recorded.

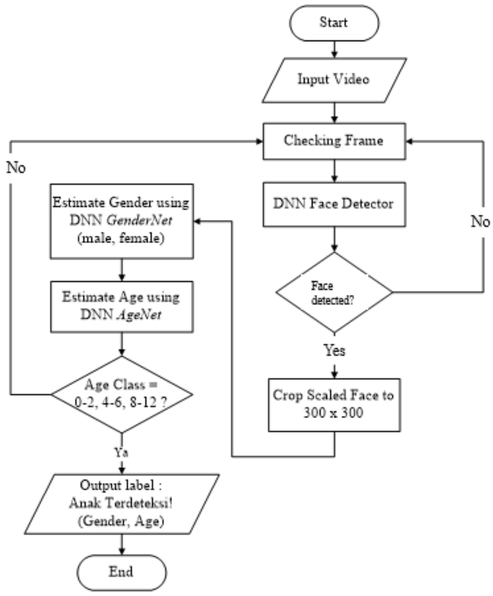

Figure 1. Flowchart of Child Presence Detection System

If the age estimation output falls within 0–2, 4–6, or 8–12 years, the system assigns the label "Child Detected" to the detected object. This pipeline supports real-time processing, optimized for low-power edge devices. The detection and classification results are recorded for further evaluation using statistical and sensitivity metrics.

B. Dataset Building

This study employs two types of data sources. First, the Adience Benchmark Database is used for pretraining and testing the age and gender estimation models. This publicly available dataset comprises 26,580 unconstrained face images from 2,284 subjects, annotated with eight age categories and binary gender labels. The dataset reflects real-world conditions with various poses, occlusions, and lighting inconsistencies.

Second, experimental video data is collected using the Canon EOS M50 camera. Each video lasts 3 seconds, producing 72 frames per scenario at 24 fps. Experimental parameters include five distances (100–500 cm), five camera heights (70–150 cm), and two lighting conditions (indoor: 110–352 lux, outdoor: 8410–9440 lux). These variations simulate diverse real-world monitoring scenarios. Preprocessing is applied using OpenCV's blobFromImage() function, including resizing to 227×227 pixels, mean subtraction, channel swapping (BGR to RGB), and pixel scaling, aligning with model input specifications.

C. Training Method

The system integrates pretrained models AgeNet and GenderNet, which are based on the CNN architecture introduced by Levi and Hassner [11]. These models are implemented using the Caffe framework and deployed via OpenCV's cv2.dnn module. The model architecture consists of 3 Convolutional layers, conv1 using 96 filters, 7×7 kernel, conv2 using 256 filters, 5×5 kernel, conv3 using 384 filters, 3×3 kernel, each followed by 3×3 max pooling with stride 2, 2 Fully-connected layers 512 neurons each, Softmax classifier to predicts one of eight age groups or two gender classes. The activation function used is ReLU, while categorical cross-entropy serves as the loss function during original training. These pretrained models are not fine-tuned further in this study, instead, their inference capabilities are evaluated in real-world settings.

D. Testing Method

After the implementation of design and programming, detection testing is carried out to determine the system's performance. The child presence detection program is run on an Asus A455L laptop with an Intel Core i3-4030U processor. The video input is an mp4 file with a resolution of 1920x1080 4000kbps 24fps obtained from a Canon EOS M50 camera.

The parameters tested in this study include: First, the light hitting the object's face, measuring the intensity of light hitting the object's face in the test environment using a lightmeter in lux units [12]. In this research, the test environment is divided into two categories, namely indoor and outdoor. The second is the camera's height against the ground. The camera is placed in a static position using a tripod with a height of 70 cm, 90 cm, 110 cm, 130 cm, and 150 cm above ground level. Then, the last is the distance between the camera and the object.

The object's position against the camera is set as far as 100 cm, 200 cm, 300 cm, 400 cm, and 500 cm. Camera height and object distance can be described as in Figure 2.

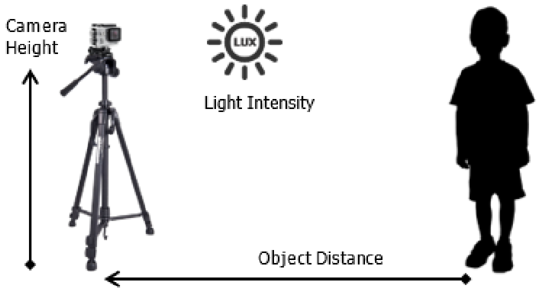

Figure 2. Testing Parameters

The test scenarios in data collection in each test environment are as follows: installing the camera device on a tripod, measuring the intensity of light hitting the face of the child object with a lightmeter, adjusting the height of the tripod to get the camera height parameter to the ground, adjusting the distance of the detected object and its position so that the face is facing the camera, recording the object as video input to the system, running the detection program on a PC/laptop device with input in the form of video from the camera, monitoring and recording the program process, and retrieving detection result data in the form of child detection success, age estimation, and gender [13]**.**

In this research, the data obtained are visual data and quantitative data. What is meant by visual data in this study is the results of screenshots when the child presence detection program is run. What is meant by quantitative data is the calculation of the success of the number of detected frames, including faces, gender estimation, and child detection for each test scenario based on predetermined parameters [14].

The data obtained from the results of testing and data collection will be analyzed to determine the system's performance and factors that can affect the system in detecting children. The performance of a classification system describes how well the system classifies data, which can be calculated using the confusion matrix. In binary classification, the input data is grouped into only one of two non-overlapping classes [15].

Sensitivity is the effectiveness of the classification system in identifying positive labels from the amount of data per class [16]. The performance analysis of the child presence detection system with the age estimation method in this study was carried out on three types of detection flows. The first is face detection, where the TP value is the actual condition when there is a face object detected correctly by the system, and the TN value is the actual condition when the system detects a non-face object; the second is gender estimation, where the TP value is the actual condition when the object's gender is detected correctly by the system and the FN value is the actual condition when the object's gender is detected incorrectly by the system. The third is child detection, where the TP value is the actual condition when a child object is detected correctly by the system, and the FN value is the actual condition when a child object is not detected by the system [17].

Meanwhile, to find out which factors can affect the system in detecting the presence of children using statistical tests. The method used is the two-way anova statistical method. This method was chosen because it can be seen which parameters significantly affect the number of successful child detections, whether light intensity, camera height to ground level, or object distance to the camera [18]—data processing with the statistical test method of the test results using the MINITAB® 18.1 software application tool.

III. RESULT AND DISCUSSION

Furthermore, data was collected for further analysis according to the specified parameters. Data collection was carried out on a sample of 3 male child objects. Each child tested variations in camera height of 70-150 cm and object distance of 100-500 cm at each test location so that the following data and test results were obtained.

A. Testing Data

In testing the system, visualization is carried out using the Python programming language. So, from the visualization, the detection results of objects considered children by the system can easily be seen. The system will give a green box to the human face object, then predict the gender and age of the object accompanied by a yellow detection confidence value. If the age estimation result detected by the system is the age group 0-2, 4-6, or 8-12, an alert will appear as a red "Child detected!" label. However, if the age estimation result detected is the age group of 15-20, 25-32, 38-43, 48-53, or 60-100, no alert will appear on the screen

B. Data Capture Results

Each video run based on the parameter scenario in this test is limited to 3 seconds. With a video input resolution of 1920x1080 4000kbps 24fps, the total frames detected for each video were 24fps x 3s = 72 frames. The data collection scenario is carried out by the testing parameters in this study, namely light intensity, camera height, and object distance.

C. Indoor Data Collection Results

Data collection was carried out indoors in an empty house with the intensity of light hitting the face of the child object in the range of 110-352 lux, as shown in Figure 3. Then, the light comes from the sunlight's reflection entering the test location [19]. Tests were carried out using variations in camera height 70-150 cm and object distance 100-500 cm.

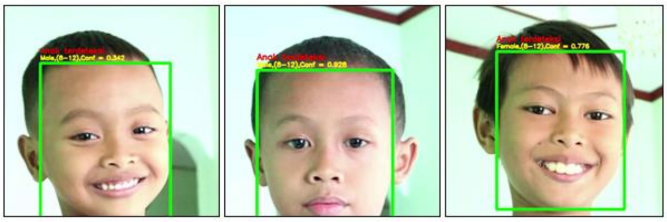

Figure 3. Testing on Child Objects with 110-352 lux

From the number of successful detections, the sensitivity value is then calculated so that the face detection sensitivity data is obtained in Figure 4, gender estimation sensitivity in Figure 5, and child detection sensitivity in Figure 6.

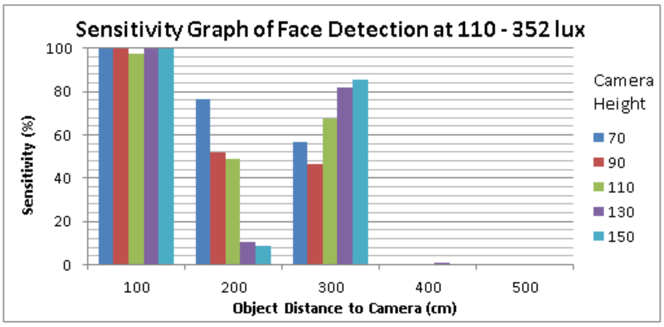

Figure 4. Comparison of Face Detection Sensitivity Value at 110-352 lux

It is known that the best face detection sensitivity value is at an object distance of 100 cm, both at a camera height of 70 cm, 90 cm, 110 cm, 130 cm, and 150 cm. Then, it can be seen that when the distance of the object to the camera shifts to 200 cm and 300 cm, there is a decrease in the performance of face detection in the system. The sensitivity value at a distance of 200 cm is smaller than that at a distance of 300 cm because the room door blocks the reflection of sunlight on the child's face. In the camera height test, the best face detection sensitivity value with an object distance of 200 cm and 300 cm is 70 cm and 150 cm, respectively. So, it is unknown which camera height significantly affects the face detection results. Meanwhile, the face detection sensitivity value is close to zero when the distance between the object and the camera is 400 cm and 500 cm. This is because the system cannot detect faces if the distance between the object and the camera is too far.

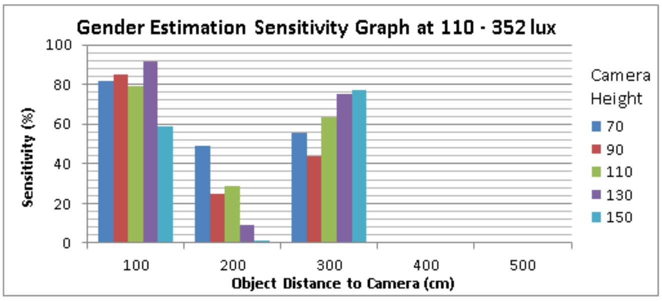

Figure 5. Comparison of Gender Estimation Sensitivity Values at 110-352 lux

The graph in Figure 5 shows that the sensitivity value of sex estimation is best at an object distance of 100 cm with a camera height of 70 cm, 90 cm, 110 cm, 130 cm, and 150 cm. Then, it can be seen that when the object's distance from the camera shifts to 200 cm and 300 cm, there is a decrease in the performance of sex estimation in the system. The sensitivity value at a distance of 200 cm is smaller than the distance of 300 cm due to the influence of the decrease in face detection performance in the previous flow. In the camera height test, the best sex estimation sensitivity values with object distances of 100 cm, 200 cm, and 300 cm are 130 cm, 70 cm, and 150 cm, respectively. Therefore, it is unknown which camera height significantly affects the performance of sex estimation. Meanwhile, when the distance between the object and the camera is 400 cm and 500 cm, the sensitivity value for detecting the presence of children is 0%.

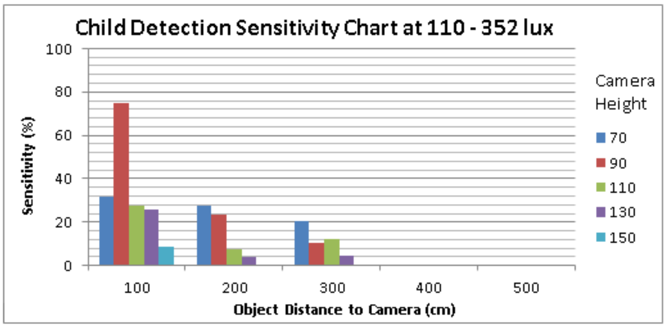

Figure 6. Comparison of Child Detection Sensitivity Value at 110-352 lux

The graph in Figure 6 shows that the best performance of child presence detection in the system is when testing a camera height of 90 cm and an object distance to the camera as far as 100 cm. Meanwhile, when the object distance is 200 cm and 300 cm, the sensitivity value of child presence detection is best when the camera height is 70 cm. So, it is unknown which camera height significantly affects the results of gender estimation.

D. Outdoor Data Collection Results

Data collection was carried out in the area around the house under construction with the intensity of light hitting the face of the child object in the range of 8410-9440 lux, as in Figure 7. Then, in this test, the object gets light from direct sunlight. Tests were carried out using variations in camera height 70-150 cm and object distance 100-500 cm.

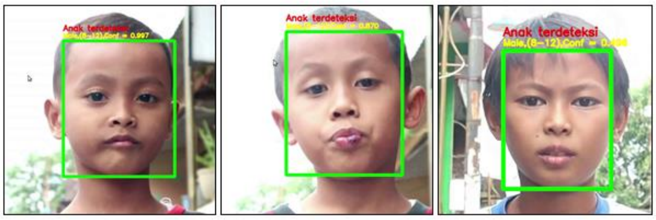

Figure 7. Testing on Child Objects with 8410-9440 lux

From the number of successful detections, the sensitivity value is then calculated so that the face detection sensitivity data in Figure 8, gender estimation sensitivity in Figure 9, and child detection sensitivity in Figure 10.

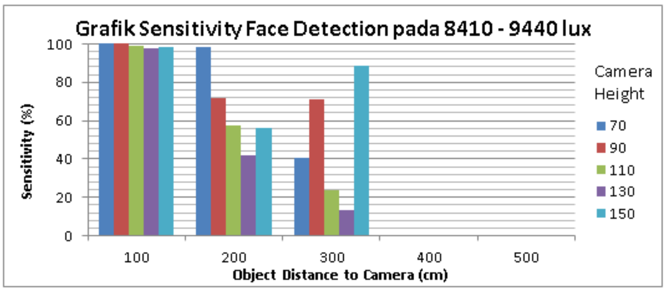

Figure 8. Comparison of Face Detection Sensitivity Value at 8410-9440 lux

It is known that the best face detection sensitivity value is at an object distance of 100 cm with a camera height of 70 cm, 90 cm, 110 cm, 130 cm, and 150 cm. Then, it can be seen that when the object's distance from the camera shifts to 200 cm and 300 cm, there is a linear decrease in the performance of face detection in the system. This is because the further the distance of the object, the more difficult it is for the system to detect faces. In the camera height test, the best face detection sensitivity value with an object distance of 200 cm and 300 cm is 70 cm and 150 cm, respectively. So, it is unknown which camera height significantly affects the face detection results. Meanwhile, the face detection sensitivity value is close to zero when the distance between the object and the camera is 400 cm and 500 cm. This is because the system cannot detect faces if the distance between the object and the camera is too far.

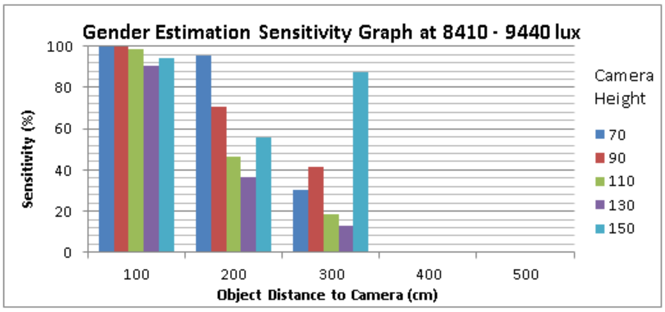

Figure 9. Comparison of Gender Estimation Sensitivity Value at 8410-9440 lux

It is known that the sensitivity value of sex estimation is best at an object distance of 100 cm with a camera height of 70 cm, 90 cm, 110 cm, 130 cm, and 150 cm. It can be seen that when the object's distance from the camera shifts to 200 cm and 300 cm, there is a linear decrease in the performance of sex estimation in the system. This is because the further the distance of the object, the more difficult it is for the system to estimate the gender. In the camera height test, the best face detection sensitivity value with an object distance of 200 cm and 300 cm is 70 cm and 150 cm, respectively. So, it is unknown which camera height significantly affects the results of gender estimation. Meanwhile, when the distance between the object and the camera is 400 cm and 500 cm, the sensitivity value of gender estimation is 0%.

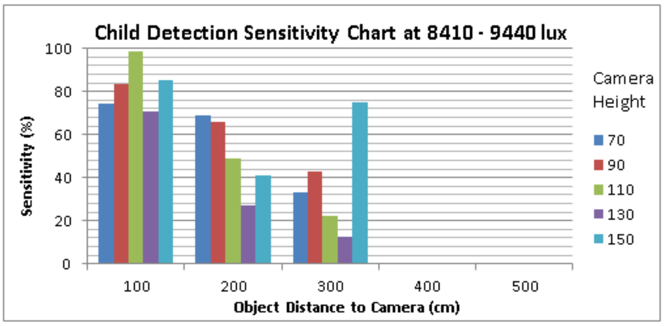

Figure 10. Comparison of Child Detection Sensitivity Value at 8410-9440 lux

It is known that the sensitivity value of child presence detection is best when the object distance is 100 cm with a camera height of 70-150 cm. Then, it can be seen that when the object's distance from the camera shifts to 200 cm and 300 cm, there is a linear decrease in detecting children's presence in the system. This is because the further the distance of the object, the more difficult it is for the system to classify the age of the child. In testing the camera height, the sensitivity value of detecting the presence of children is best with an object distance of 200 cm and 300 cm, namely 70 cm and 150 cm, respectively. So, it is unknown which camera height significantly affects the gender estimation results. Meanwhile, when the distance between the object and the camera is 400 cm and 500 cm, the face detection sensitivity value is 0%.

E. Data Analysis

From the test data obtained and the initial description and analysis based on the graph in the previous section, statistical calculations can be carried out to prove and strengthen what factors affect the results of child detection in the system. Data processing is carried out using the two-way anova statistical method using MINITAB® 18.1 software.

Hypothesis testing on the test data of the results of detecting the presence of children in this study was carried out to determine whether there was an effect of testing parameters, namely indoor and outdoor light intensity, camera height, and camera distance, on the number of successful child detections. The first step to perform hypothesis testing is to determine H0 and H1 ($H_{Alternative}$). H0 in this test is that there is no significant influence between the test parameters on the number of successful child detections. While H1 ($H_{Alternative}$) in this test is that there is a significant influence between the test parameters on the number of successful child detections. The data confidence level in this test is 95%, so the α value is 100%-95% = 5% (0.05). In using MINITAB® 18.1 software with the two-way anova statistical method, the Analysis of Variance results are obtained as shown in Table 1.

TABEL 1
MINTAB OUTPUT: ANALYSIS OF VARIANCE

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Light Intensity	1	10135	10135,3	41,34	0,000
Camera Distance	4	2080	520,1	2,12	0,081
Object Distance	4	36673	9168,3	37,39	0,000
Error	140	34326	245,2		
Lack-of-Fit	40	16835	420,9	2,41	0,000
Pure Error	100	17491	174,9		
Total	149	83215			

Table 1 shows that the p-value for the camera height factor has a value of 0.081 > α: 0,05. This explains that H0 is accepted or there is no significant influence between camera height and the number of successful child detections. At the same time, the p-value for the light intensity factor and object distance has a value of 0.000 < α: 0,05. This explains that H0 is rejected and H1 ($H_{Alternative}$) is accepted, or there is a significant influence between light intensity and object distance on the number of successful child detections. Then, in the same way, it was tested again to find out whether there was also an interaction effect between the factors of light intensity and object distance. The Analysis of Variance results obtained are shown in Table 2.

TABLE 2
MINITAB OUTPUT: ANALYSIS OF VARIANCE PADA INTERACTION FACTOR

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Light Intensity	1	10135	10135,3	49,29	0,000
Object Distance	4	36673	9168,3	44,58	0,000
Light Intensity*Object Distance	4	7616	1904,0	9,26	0,000
Error	140	28790	205,6		
Total	149	83215			

Table 2 shows that the p-value for the interaction between the factors of light intensity and object distance has a value of 0.000 < α: 0,05. This explains that H0 is rejected and H1 ($H_{Alternative}$) is accepted, or there is a significant influence between the interaction of light intensity and object distance on the number of successful child detections. To find out which light intensity, object distance, and the interaction of the two factors provide the most optimal results on the number of successful child detections, it is necessary to conduct a Tukey further test on the three parameters, as shown in Tables 3, 4, and 5.

TABLE 3
TUKEY PAIRWISE COMPARISONS (LIGHT INTENSITY)

Tukey Pairwise Comparisons: Light Intensity

Grouping Information Using the Tukey Method and 95% Confidence

Light Intensity	N	Mean	Grouping	
8410 - 9440	75	24,48	A	
110 - 352	75	8,04		B

Means that do not share a letter are significantly different.

As seen in Table 3, the significance level of the light intensity factor is divided into two groups, namely groups A and B. Testing with light intensity in the range of 8410-9440 lux is a member of group A, which results in a higher average number of successful child detections than group B, which is light intensity in the range of 110-352 lux.

TABLE 4
TUKEY PAIRWISE COMPARISONS (OBJECT DISTANCE)

Tukey Pairwise Comparisons: Object Distance

Grouping Information Using the Tukey Method and 95% Confidence

Object Distance	N	Mean	Grouping		
100	30	41,8000	A		
200	30	22,6667		B	
300	30	16,8333		B	
500	30	0,0000			C
400	30	0,0000			C

Means that do not share a letter are significantly different.

As seen in Table 4, the significance level of the object distance factor to the camera is divided into three groups: A, B, and C. The object distance of 100 cm included in group A results in the highest average number of successful child detections. Object distances of 200 cm and 300 cm included in Group B resulted in an average number of successful child detections lower than Group A but higher than Group C. Object distances of 400 cm and 500 cm included in Group C resulted in the lowest average number of successful child detections. In addition to the grouping table, the significance of the object distance factor can also be shown through the main effect plot graph in Figure 11.

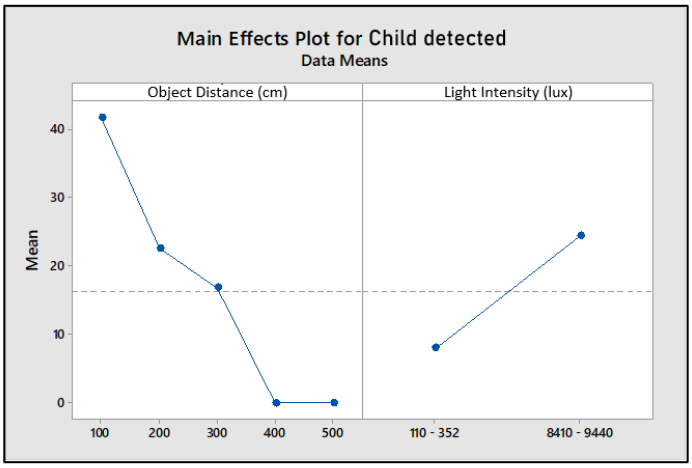

Figure 11. Main Effect Plot of Object Distance and Light Intensity Parameters

The main effect plot graph in Figure 11 shows that testing with object distances of 100 cm, 200 cm, and 300 cm resulted in an above-average number of successfully detected children. Testing with object distances of 400 cm and 500 cm results in the number of successful children detected below average. It can also be seen that testing with outdoor light intensity with a range of 8410-9440 lux resulted in a higher number of successfully detected children than indoor light intensity with a range of 110-352 lux.

TABLE 5
TUKEY PAIRWISE COMPARISONS (LIGHT INTENSITY*OBJECT DISTANCE)

Tukey Pairwise Comparisons: Light Intensity*Object Distance

Grouping Information Using the Tukey Method and 95% Confidence

Light Intensity (lux)*Object Distance(cm)	N	Mean	Grouping			
8410 - 9440 100	15	59,3333	A			
8410 - 9440 200	15	36,2667		B		
8410 - 9440 300	15	26,8000		B		
110 - 352 100	15	24,2667		B	C	
110 - 352 200	15	9,0667			C	D
110 - 352 300	15	6,8667				D
110 - 352 500	15	0,0000				D
8410 - 9440 500	15	0,0000				D
110 - 352 400	15	0,0000				D
8410 - 9440 400	15	0,0000				D

Means that do not share a letter are significantly different.

As seen in Table 5, the significance level of the interaction parameter of light intensity with the distance of the object to the camera is divided into four groups, namely groups A, B, C, and D. The interaction of light intensity 8410-9440 lux with an object distance of 100 cm included in group A produces the highest average number of successful child detections. The interaction of light intensity 8410-9440 lux with an object distance of 200 cm and 300 cm and light intensity 110-352 lux with an object distance of 100 cm is included in group B. The interaction of light intensity 110-352 lux with object distances of 100 cm and 200 cm is included in group C. The interaction of light intensity 110-352 lux with object distances of 200 cm and 300 cm and light intensity both 110-352 lux and 8410-9440 lux with object distances of 400 cm and 500 cm are included in group D resulting in the lowest average number of successful child detections. In addition to the grouping table, the significance of the interaction parameter of light intensity with object distance can also be shown through the interaction plot graph in Figure 12.

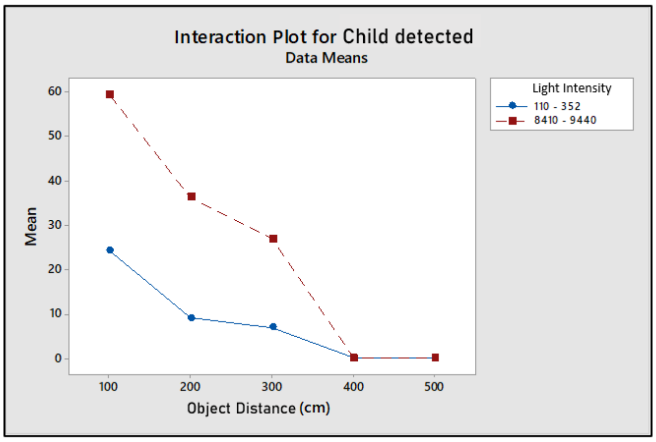

Figure 12. Interaction Plot of Light Intensity*Object Distance Parameters

In the interaction plot graph in Figure 12, it can be seen that the interaction parameters of light intensity with object distance that produce the best number of successfully detected children are outdoor light intensity with object distances of 100 cm, 200 cm, 300 cm, indoor light intensity with object distances of 100 cm, 200 cm, 300 cm, light intensity both indoors and outdoors with object distances of 400 cm and 500 cm, respectively.
Based on the discussion and analysis of the previous test data, it is known that the light intensity and the distance of the object to the camera significantly affect the sensitivity value in the three detection flows. The sensitivity value at a certain distance is shown in Table 6.

TABLE 6
COMPARISON OF SENSITIVITY VALUES AT SPECIFIC DISTANCES

Object Distance (cm)	110-352 lux			8410-9440 lux		
	Face Detection Sensitivity (%)	Gender Estimation Sensitivity (%)	Child Detection Sensitivity (%)	Face Detection Sensitivity (%)	Gender Estimation Sensitivity (%)	Child Detection Sensitivity (%)
100	99,50	77,78	32,24	99,01	96,63	84,72
200	41,17	23,31	12,80	65,58	62,20	51,39
300	65,67	60,62	9,33	50,79	40,97	39,88
400	0,10	0	0	0	0	0
500	0	0	0	0	0	0

The detection of the presence of children with the age estimation method in this study is influenced by several conditions, such as Differences in light intensity. When the object gets light from direct sunlight outside the room, the system can detect the presence of children with good results. In tests conducted by utilizing reflected sunlight that enters the room, the system makes it difficult to detect the presence of children. The reflection of sunlight on objects around the room can affect the system's sensitivity to detect the presence of children. Then, the object's distance to the camera is influenced. In testing the distance of the object to the camera, there is a linear decrease in the performance of detecting the presence of children in the system. The further the object's distance to the camera, the more difficult it is for the system to classify the child's age. This is because the effect of low focal length on the camera causes objects near the camera to appear larger and objects farther from the camera to appear smaller.

Furthermore, the ability of the system to detect faces, if there is a decrease in face detection performance, the sensitivity of detecting the presence of children also decreases. This is because the system will only estimate age and gender if a human face object is detected in the frame. Finally, Prediction flickering. The processing of matrix information to detect the presence of children is carried out on each frame of an input video. The system considers each frame as a different object. This causes the detection label to be unstable (prediction flickering).

IV. CONSLUSION

This research aimed to develop and evaluate a child presence detection system using an age estimation method based on facial features, implemented through computer vision techniques. The system successfully achieved the objectives outlined in the early stages of the study, particularly in demonstrating that children's presence can be effectively detected through face recognition and age classification using external pretrained models, AgeNet and GenderNet integrated via the OpenCV DNN module. Based on statistical evaluation using the two-way ANOVA method with a 95% confidence level ($\alpha = 0.05$), it was found that object distance and light intensity significantly affected the system's detection performance (p-value = 0.000 < 0.05), while camera height did not have a statistically significant effect (p-value = 0.081 > 0.05). The system showed optimal performance at a distance of 100 cm and under natural daylight with light intensity levels between 8320–9440 lux, achieving a child detection sensitivity of 84.72%. In comparison, face detection reached 99.50%, and age-gender estimation attained 96.63% sensitivity, indicating that the proposed method performs reliably under well-lit conditions and short-range camera placement.

A key contribution of this research lies not in the development of novel detection algorithms, but in the systematic implementation, evaluation, and contextual testing of an integrated DNN-based age estimation system for child safety applications. The use of third-party pretrained models such as those based on the Levi and Hassner architecture played a central role in the success of this system and is duly acknowledged as a critical component of the overall framework. Future work is recommended to focus on integrating the current detection model with real-time alert systems and IoT-based surveillance architectures. This would enable proactive safety measures, such as live notifications or automated warnings, when a child enters a high-risk area. Additionally, expanding testing to include diverse environmental conditions, face orientations, and larger sample sizes would enhance the robustness and generalizability of the system in real-world deployments.

REFERENCES

[1] A. Deshmukh, “Utilizing Visual-Spatial Perception to Detect Drowning-induced Suffocation in Swimming Pools,” *2024 IEEE 15th Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2024*, pp. 164–168, 2024, doi: 10.1109/UEMCON62879.2024.10754739.

[2] A. Naik, S. Bhat, and D. Naik, “A Vision Based Approach for Real Time Detection of Proximity Hazards around Open Wells: An Application in Open Well Safety Monitoring,” *2024 IEEE Conf. Eng. Informatics, ICEI 2024*, pp. 1–6, 2024, doi: 10.1109/ICEI64305.2024.10912307.

[3] F. A. Khan and A. Dey, “Towards Enhancing Child Safety: A Deep Learning Approach to Detect Child Safe and Unsafe Objects,” *Proc. 2024 IEEE Int. Women Eng. Conf. Electr. Comput. Eng. WIECON-ECE 2024*, pp. 123–128, 2024, doi: 10.1109/WIECON-ECE64149.2024.10914992.

[4] M. M. Islam and J. H. Baek, “Deep learning based real age and gender estimation from unconstrained face image towards smart store customer relationship management,” *Appl. Sci.*, vol. 11, no. 10, 2021, doi: 10.3390/app11104549.

[5] D. B. Kinthadi, A. Burugu, A. Rumandla, and S. Sujana, “Prediction of Human Age and Gender Using Deep Learning For Smart Security Systems,” *2nd IEEE Int. Conf. Distrib. Comput. Electr. Circuits Electron. ICDCECE 2023*, pp. 1–6, 2023, doi: 10.1109/ICDCECE57866.2023.10151119.

[6] Y. Deng, S. Teng, L. Fei, W. Zhang, and I. Rida, “A Multifeature Learning and Fusion Network for Facial Age Estimation,” *Sensors (Basel).*, vol. 21, no. 13, pp. 1–15, 2021, doi: 10.3390/s21134597.

[7] M. Chao, R. Stoleru, L. Jin, S. Yao, M. Maurice, and R. Blalock, “AMVP : A daptive CNN-based M ultitask V ideo P rocessing on Mobile Stream Processing Platforms”.

[8] A. Karatzas and I. Anagnostopoulos, “Balancing Throughput and Fair Execution of Multi-DNN Workloads on Heterogeneous Embedded Devices,” *IEEE Trans. Emerg. Top. Comput.*, pp. 1–14, 2024, doi: 10.1109/TETC.2024.3407055.

[9] S. Sun, Haomiao and Pan, Hongyu and Han, Hu and Shan, “Deep Conditional Distribution Learning for Age Estimation,” *IEEE Trans. Inf. Forensics Secur.*, pp. 4679–4690, 2021, doi: 10.1109/TIFS.2021.3114066.

[10] Z. A. T. Ahmed *et al.*, “Facial Features Detection System to Identify Children with Autism Spectrum Disorder: Deep Learning Models,” *Comput. Math. Methods Med.*, vol. 2022, 2022, doi: 10.1155/2022/3941049.

[11] T. Levi, Gil and Hassner, “Age and Gender Classification Using Convolutional Neural Networks,” *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 34–42, Dec. 2015, doi: 10.1109/TIFS.2014.2359646.

[12] D. Agung Al Ayubi *et al.*, “Pendeteksi Wajah Secara Real Time pada 2 Degree of Freedom (DOF) Kepala Robot Menggunakan Deep Integral Image Cascade,” *CYCLOTRON*, vol. 3, no. 1, Feb. 2020, doi: 10.30651/CL.V3I1.4306.

[13] S. S. Hidayat, D. Rahmawati, M. C. Ardi Prabowo, L. Triyono, and F. T. Putri, “Determining the Rice Seeds Quality Using Convolutional Neural Network,” *Int. J. Informatics Vis.*, vol. 7, no. 2, pp. 527–534, 2023, doi: 10.30630/joiv.7.2.1175.

[14] P. Chyan and Y. Kasmara, “Sistem Monitoring dan Deteksi Stres Pada Anak Berbasis Wearable Device,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 5, pp. 943–949, Oct. 2021, doi: 10.29207/resti.v5i5.3503.

[15] R. I. Tiyar and D. H. Fudholi, “Kajian Pengaruh Dataset dan Bias Dataset terhadap Performa Akurasi Deteksi Objek,” *Petir*, vol. 14, no. 2, pp. 258–268, 2021, doi: 10.33322/petir.v14i2.1350.

[16] N. Nafi’iyah *et al.*, “System of gender identification and age estimation from radiography: a review,” *Int. J. Electr. Comput. Eng.*, vol. 13, no. 5, pp. 5491–5500, Oct. 2023, doi: 10.11591/IJECE.V13I5.PP5491-5500.

[17] S. N. D. Chua, S. F. Lim, S. N. Lai, and T. K. Chang, “Development of a Child Detection System with Artificial Intelligence Using Object Detection Method,” *J. Electr. Eng. Technol.*, vol. 14, no. 6, pp. 2523–2529, Nov. 2019, doi: 10.1007/S42835-019-00255-1/METRICS.

[18] H. I. Assaad, Y. Hou, L. Zhou, R. J. Carroll, and G. Wu, “Rapid publication-ready MS-Word tables for two-way ANOVA,” *Springerplus*, vol. 4, no. 1, pp. 1–9, Jan. 2015, doi: 10.1186/S40064-015-0795-Z/TABLES/4.

[19] M. C. A. Prabowo and S. S. Hidayat, “Edge Detection Technique for Rice Quality Analysis Using Digital Image Processing,” *AIP Conf. Proc.*, vol. 2431, no. 1, 2023, doi: 10.1063/5.0117510.