International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 16, No.1 (Sep-2024)

http://dx.doi.org/10.12785/ijcds/160199

An Automatic Approach to Detect Girl Child Trafficking

Dhrubajit Kakati[1], Rajarshi Seal[1], Tania Sarkar[1] and Ranjan Maity[1]

[1]*Department of CSE, Central Institute of Technology Kokrajhar, Kokrajhar, India*

Received 8 Sep. 2023, Revised 5 Jun. 2024, Accepted 6 Jun. 2024, Published 19 Sep. 2024

Abstract: Girl child trafficking has become a matter of serious concern for human society. There are different manual approaches to stop and prevent it. However, these approaches need a huge amount of manual interventions. Consequently, there is a necessity to develop an automatic approach for detecting the incidents of girl child trafficking. In this work, we proposed a two-stage computational model for automatic girl child trafficking by analyzing images. Due to the unavailability of girl child trafficking images, we constructed a data set having one thousand four hundred ninety-six data. After careful observations, we decided to consider three features - age, emotion, and gender. Using these three features we developed our proposed computational model. In the first stage, the *ResNet 50* deep neural network was used to determine the three feature values from an image. It was observed that these three models can perform the gender, age, and emotions with a testing accuracy of 80.23%, 76.29%, and 85.73%, respectively. In the next level, a Support Vector Machine (SVM) was used to determine whether there is a possibility of girl child trafficking or not. A K-fold cross-validation technique with K= 6 was used to avoid the overfitting problems. It has been observed our proposed model can detect girl child trafficking with an accuracy of 93.13%. The high accuracy observed in our study indicates the candidatures of our model for real-time child trafficking.
Keywords: girl child trafficking, deep learning, machine learning, image processing, Support Vector Machine.

1. Introduction

According to the UN, child trafficking encompasses the *"recruitment, transportation, transfer, harboring, and receipt kidnapping of a child for slavery, forced labor, and exploitation"*. Once ensnared in this sinister web, trafficked children are subjected to perilous and exploitative environments. This insidious practice amounts to a contemporary form of slavery, where young lives are coerced into hazardous and exploitative labor conditions. Perpetrators employ a range of tactics, including promises of a brighter future, intimidation, threats, and violence, to ensnare their victims. Child trafficking is a grave issue, inflicting profound harm on children's physical and mental well-being. It demands urgent attention and concerted efforts to combat this egregious violation of human rights.

Illicit trade is a serious issue that can only be stopped with strict correction measures. The first step towards combating trafficking is to coordinate all agencies and authorities. It is imperative that all departments work together and create an effective action plan, otherwise, the current situation may not improve. This will allow policymakers and investigation agencies worldwide to develop strong foundations for effective policy enforcement. As a result, an action plan can be formulated for an immediate response to such crimes.

Nowadays, monitoring of different public places – railway stations, bus stands, and hotels are monitored through video surveillance. However, this requires a lot of resources including manpower. Therefore, the necessity of an automatic child trafficking system requires it to be developed and employed. By automatically scanning large amounts of data from surveillance video, it may be possible to identify potential victims of trafficking and intervene before they are harmed. To the best of our knowledge, to date, there is no such automatic tool, even no comprehensive data set is found in the modern popular data repositories. The main objective of this work is -

- Development of an automatic model for girl child trafficking detection by -
  - Analyzing the age of a person.
  - Identifying the gender.
  - Analyzing emotional state.

In this work, we developed a comprehensive dataset of size one thousand four hundred. Three different models – age detection, emotion detection, and gender detection were proposed. The age detection model is suitable for detecting whether in an image there is a child or not, while the gender model detects whether the child is a girl child or not. The emotion detection model is used to identify their emotion levels – fear and sadness. All these models used Resnet 50 convolutional neural network. After the detection of the three parameters – age, gender, and emotional level, they were used to predict whether there was an incident of girl child trafficking or not. The final model works based

E-mail address: u20cse1032@cit.ac.in, u20cse1054@cit.ac.in, t.sarkar@cit.ac.in, r.maity@cit.ac.in

on the Support Vector Machine classifier. An accuracy of 93.13% was observed in the proposed model. The main contributions of our work are -

- Development of dataset for girl child trafficking.
- Development of a computational model for age prediction.
- Development of a computational model for gender prediction.
- Development of a computational model for emotion prediction.
- By combining the above three models of age, gender, and emotion - the development of a model for girl child trafficking detection.

In the development of the child trafficking prevention model, we recognize that it is not practical or accurate to assume that the victims of trafficking can be identified definitively only through image analysis. Therefore, we adopt a probabilistic approach in the proposed model, focusing on determining the potential for victims of trafficking based on image characteristics and context information. By abandoning absolute certainty, the proposed model aims to provide a probability indicator of whether images contain potential indicators of child trafficking. This probabilistic view allows us to avoid making final judgments based solely on images and emphasizes the importance of using technology as an additional tool to support and strengthen human efforts in the fight against human trafficking. Through this approach, we are committed to creating a judicious and responsible model that can contribute significantly to the early detection and prevention of trafficking of girl child, while respecting the complexity and uncertainties inherent in this crucial social issue.

2. Related Work

The proposed work has three fundamental pillars: age, gender, and emotion. These pillars serve as crucial components in the analysis of human behavior and psychological patterns. Through the utilization of diverse methodologies and algorithms, accurate determination and examination of age, gender, and emotion are made possible. The references cited in this study serve as valuable sources that underpin the methodologies and algorithms employed. In the following, we briefly discussed the existing approach of age, gender, and emotion detection.

Our literature review begins by acknowledging the limited existing research in the chosen area of investigation. As pioneers in this field, we embrace the challenge of charting new territories and conceptualizing fresh perspectives. While the scarcity of prior studies poses inherent difficulties, it also grants us the unique opportunity to present a pioneering narrative that incorporates relevant insights from tangential disciplines. Our synthesis of these diverse sources aims to provide a robust theoretical framework, fueling the trajectory of our research. Through diligent analysis and a quest for understanding, we aspire to uncover unexplored dimensions, laying the groundwork for future investigations and pushing the boundaries of knowledge in this emerging domain.

We acknowledge the lack of a well-established knowledge base in the area of interest. However, this void encourages us to adopt a creative and open-minded approach to the literature review. By delving into adjacent fields, we extract valuable concepts and methodologies that can enrich our study. Our exploration of analogous topics allows us to draw insightful parallels and extrapolate potential implications for our own research. While the absence of existing work poses challenges, it also provides an exciting prospect for innovation and discovery. Through meticulous analysis and a collaborative spirit, we are poised to contribute a pioneering perspective that can shape the trajectory of future research in this domain.

A. Age Detection

In literature, the exploration of diverse methodologies for age estimation is emphasized. Abirami et al. [1] advocate for a methodology involving real-time video capture, face region detection using Haar Cascades, and age prediction through a Caffenet deep CNN framework, achieving an overall accuracy of 68.89%. Antipova et al. [2] reported state-of-the-art results using convolutional neural networks (CNNs) for automatic age estimation from facial images, triumphing in the 2016 ChaLearn Apparent Age Estimation Challenge. Al-Shannaq and Elrefaei [3] present an exhaustive evaluation of various age estimation approaches, including handcrafted-based and deep learning-based models, and multi-feature fusion, detailing their respective strengths and weaknesses. Janahiraman and Subramaniam [4] delve into gender classification with Convolutional Neural Network (CNN) based Deep Learning architectures, comparing VGG-16, ResNet-50, and MobileNet, with VGG-16 demonstrating superior accuracy. Levi and Hassner [5] propose a simple convolutional net architecture applicable even with limited learning data. Nga et al. [6] present a transfer learning pipeline for age and gender prediction using facial images, leveraging pre-trained ImageNet models for superior performance. Teh and Taylor [7] address the estimation of apparent age, focusing on predicting someone's perceived age rather than their actual age. Their approach involves a CNN module capturing spatial relationships and a relational network enhancing age estimation accuracy, contributing promising results to the understudied field. The subsequent section details related work specifically associated with gender detection.

B. Gender Detection

Gender detection by scanning an image has also been explored over the years. Khan et al. [8] reported a thorough review of state-of-the-art gender classification techniques by analyzing their strengths, weaknesses, and standard datasets. It offers valuable insights for researchers in this field, considering future directions based on existing limitations. Dhomne et al. [9] addressed the limitations of existing systems for gender recognition in online social media by proposing a Deep Convolutional Neural Network (D-CNN) based on the VGGNet architecture, achieving improved performance with limited training data. The application of data mining and Delaunay triangulation for gender detection through frontal facial images and the used Classification algorithms - Functional Trees, AdaBoost, and J48 were explored in [10]. Abirami et al. [1] proposed a methodology that involves capturing video in real-time, detecting the face region in the image using Haar Cascades, and using a Caffenet deep CNN framework for gender prediction which gave an overall accuracy of 74.55%. Besides age and gender detection, emotion detection models also have a great role to play in girl-child trafficking. In the following, we briefly summarized the existing works on human emotion detection.

C. Emotion Detection

Researchers hold varying perspectives on the number of emotional states in humans, with a notable categorization of seven emotional states detailed in [11]: anger, disgust, fear, happiness, sadness, surprise, and contempt. Machine learning and natural language processing play pivotal roles in human emotion detection. Jaiswal et al. [12] propose an artificial intelligence system utilizing a convolutional neural network-based deep learning architecture for emotion detection from facial expressions. This three-step process involves face detection, feature extraction, and emotion classification, achieving an accuracy of 70.14% and 98.65% on the FERC-2013 and JAFFE datasets, respectively. Zheng et al. [13] introduce a novel method for emotion recognition from non-frontal facial images, employing regional covariance matrix representation and a discriminant analysis approach without requiring face alignment or landmark localization, showcasing its effectiveness through extensive experiments on a generated database. Recent advancements highlight the success of deep neural networks in automatic affect recognition, integrating auditory and visual modalities. Tzirakis et al. [14] propose an emotion recognition system using a combination of convolutional neural networks and deep residual networks, augmented by long short-term memory networks, outperforming traditional approaches on the RECOLA database. Siam et al. [15] employ the Viola-Jones algorithm to detect various emotions from frontal facial images, including anger, contempt, disgust, fear, happiness, sadness, and surprise. A comprehensive emotion detection system focusing on eye and mouth expressions is detailed in [16], utilizing skin detection, eye and mouth detection, and emotion recognition from color images. Ng et al. [17] leverage transfer learning with deep Convolutional Neural Networks (CNNs) for static facial expression recognition in the Emotion Recognition in the Wild contest, achieving significant improvements through a two-stage fine-tuning process on relevant datasets, surpassing the challenge baseline with an accuracy of 48.5% (validation) and 55.6% (test).

To the best of our knowledge, we did not find any automatic approach for automatic girl child trafficking in literature. After careful analysis, we decided on the three factors - age, gender, and emotion for the proposed work. Based on these features, we developed the proposed model. As we observed a high accuracy of 93.13%, we may claim that the model may be effective for the automatic detection of girl child trafficking. It may be observed that machine learning and deep learning models are popularly used for age, gender, and emotion detection. However, the usage of them to detect girl child trafficking is still not explored to date. The age factor plays an important role in identifying potential victims, as children are the primary targets of traffickers. Analyzing age-related characteristics, such as physical appearance and developmental stages can lead to further investigations. Gender is another important factor in detecting child trafficking. It helps to recognize patterns and identify potential vulnerabilities based on gender-specific risks. For Example, girls are more vulnerable to sexual exploitation trafficking, and boys are more vulnerable to forced labor or recruitment as child soldiers. In addition, including emotion analysis can provide additional information layers. Trafficked children often show emotional distress, fear, or trauma. Using facial expression recognition or emotional analysis techniques, we can detect emotional distress indicators in images. This helps to identify potential victims who may have been forced or psychologically manipulated. Combining three characteristics – age, gender, and emotional analysis can create a more comprehensive and effective system for the automatic detection of possible child trafficking cases. It should be noted that this method is not foolproof, and may generate false positives or miss certain cases, but can still be an effective tool to help law enforcement and organizations combat this horrific crime. It can help identify victims early, collect evidence for prosecution, and ultimately save lives and protect vulnerable children from the horrors of trafficking. In the following, the materials and methods used in order to develop girl child trafficking are discussed.

3. Materials and Methods

We planned to use a machine and deep learning-based approach to develop the proposed model. These types of approaches are data-dependent and are standard data sets for age, gender, and emotion, unfortunately, we did not find any such comprehensive image data set on girl child trafficking. In the following, we reported the data-set development for the proposed model. Implementing a model for preventing girls' trafficking in children based on computer vision poses several challenges in the real world, which must be addressed in order to effectively implement it. One of the major challenges is the variability of surveillance video quality. In public places, surveillance cameras often capture low-resolution, noise, and motion noise images that affect model analysis accuracy. In addition, the distance between cameras and subjects may result in limited visibility and ambiguous facial features, making it difficult to extract meaningful information for classification. To address these challenges, preprocessing techniques - image enhancement and resolution enhancement algorithms are used to improve the quality of input images before integrating them into the model.

A. Development of Data-set

This research uses a unique method to construct test data sets to deal with the sensitive nature of human trafficking images. As it was impossible to obtain direct access to real trafficking images from law enforcement agencies or online sources, alternative methods were adopted due to privacy and ethical concerns. Test data are carefully curated from various online news platforms and websites and require anonymity to ensure data source confidentiality. These news platforms cover human trafficking incidents without explicitly providing real trafficking images, as the content is sensitive. In order to construct the test data set, images were selected from news articles about human trafficking. The images used in the dataset are contextual and represent human trafficking, without directly revealing the identity of the victims or compromising their privacy. The images were further validated by specialized experts to ensure that they accurately depicted the relevant situations relating to human trafficking without causing damage or suffering to the victims. Although this method may introduce some limitations in terms of direct access to actual trafficked images, it offers researchers a responsible and ethical solution to studying the problem while preserving the privacy and dignity of the victims. Using this anonymized dataset, we were able to rigorously evaluate the proposed computational methods for the detection of automatic girl trafficking while respecting the subject sensitivity.

We manually constructed a comprehensive data set containing images specially curated to represent potential scenarios of child trafficking in girls. This data set allowed us to address the scarcity of data publicly available on this sensitive topic and ethical sources. Afterward, we focused on three key attributes - age, gender, and emotion - as important indicators for detecting trafficking. To classify these attributes, we use a predefined model and fine-tune it to the unique characteristics of our database. ResNet-50 is a deep learning architecture known for image classification and is used to train and test models. The obtained classification results are then stored in structured CSV format, enabling easy access for further analysis and interpretation. The research aims to apply support vector machines (SVMs) to processed data. The SVMs used ResNet-50's powerful representations to create a clear separation between trafficking and non-trafficking in our data set. The adoption of SVM enabled us to achieve a high degree of accuracy in the differentiation of images containing potential trafficking indicators from those that do not. The unique combination of the manually curated dataset, attribute-based classification using predefined models, ResNet-50's deep learning capability, and SVM's robust classification capability form the pillar of our research's unique contribution.

To develop the proposed three models of age, gender, and emotion, we have used the existing data sets. The age data set was obtained from Kaggle [1]. It consists of 9,000 images divided into four age classes: Class 1 (0-20 years), Class 2 (21-30 years), Class 3 (31-49 years), and Class 4 (above 50 years). It may be noted that Class 1 with an age range from 0 - 20 years is considered a child. Kaggle gender data set having twenty-seven thousand images with two labeled classes – male and female are used for gender detection. A comprehensive Kaggle data set of thirty-five thousand images is considered for seven levels of emotional states - happy, sad, angry, neutral, surprised, fearful, and disgusted. Out of these seven emotions, the sad, angry, fearful, and disgusted emotions are considered to be the emotions associated with the case of trafficking. After these three different data sets, we need to develop a girl child trafficking data set. Altogether, one thousand four hundred images having different ages, genders, and emotion levels are collected from the internet. The source of the images includes various news articles and different photo galleries –

- picsart (https://picsart.com/)
- istockphotos (https://www.istockphoto.com/)
- gettyimages (https://www.gettyimages.in/).

Using web scraping, the images were downloaded. Python and Selenium libraries were used for that purpose. In Figure 1, four such collected images are shown. The images are intentionally masked to avoid the identity disclosure of those children. The developed data set needs to be pre-processed for the model development. The pre-processing step used by us is discussed in the following.

B. Pre-processing

Pre-processing steps for the classification of age, gender, emotion, and child trafficking images include preparing face images for accurate analysis and prediction. These steps ensure that input images are properly formatted and converted to enable effective processing. First, the facial images are detected and cut to focus exclusively on the face. This step uses the face detection algorithm to identify the face regions in the image. By isolating faces, we remove irrelevant information and improve the accuracy of the later analysis.

Then, the image of the cropped faces was resized to a consistent size, usually a square shape, to ensure uniformity throughout the sample. This step is important for maintaining the aspect ratio and enabling efficient processing. Various pre-processing techniques are used to improve the quality and interpretation of images. These include standardization to adjust the pixel values in the standard range,

[1] kaggle.com

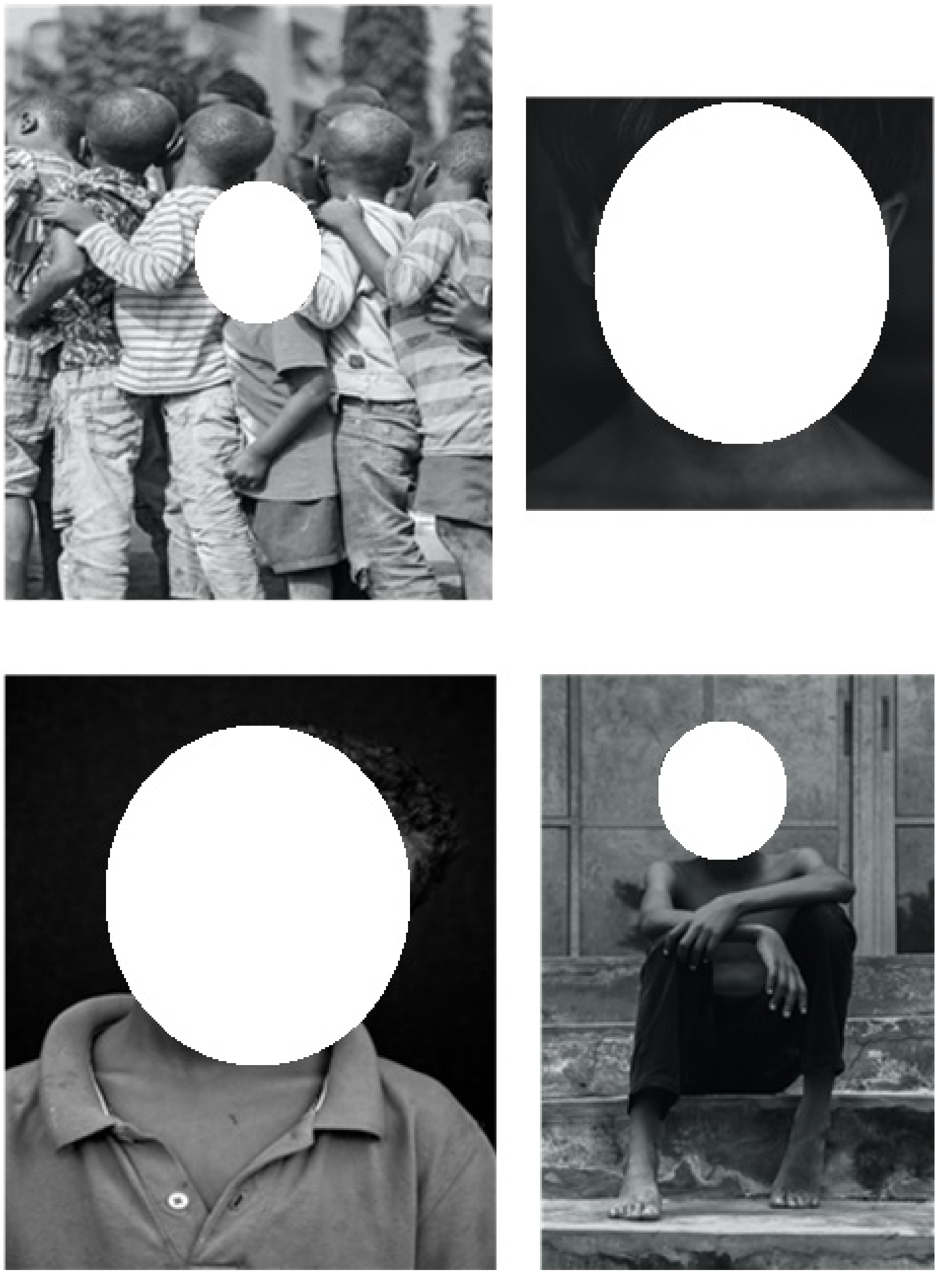


Figure 1. Sample images used in our data set.

such as scaling between 0 and 1, histogram adjustment, or contrast adjustment.

For emotion classification, performing additional pre-processing, such as facial landmark detection, may be useful to capture facial expressions accurately. This step involves identifying specific points of the face such as the eyes, nose, and mouth, which provide valuable information for emotion analysis. Overall, the pre-processing process is aimed at standardizing the size, improving the quality, and extracting the relevant facial features of the image. By carefully preparing the images, we ensured that the prediction models were optimally able to perform the task of age, gender, emotion, and child trafficking image classification. Based on the pre-processed images, we developed the proposed model for child trafficking as discussed below.

C. Proposed Model

In this section, we present four proposed models for the classification of age, gender, emotions in facial images, and girl child trafficking. To develop the models of age detection, we used the age dataset of *Kaggle* and a *ResNet50*

Convolutional Neural Network (CNN). Similarly, using the dataset of gender, and emotion we developed two different models based on ResNet50 (CNN) for gender and emotion, respectively.

ResNet-50 has garnered substantial recognition in the realms of computer vision and image classification, standing out as a prominent convolutional neural network architecture. In benchmark datasets, residual neural networks achieve higher accuracy than conventional deep neural networks. The training process in ResNet is generally faster than the other deep networks due to faster convergence. Due to the generalization, it also provides better results during testing. ResNet-50, a member of the ResNet family, short for "Residual Network," has gained renown for its exceptional performance in deep learning applications. Comprising 50 layers, it stands out as a notably deep network. A pivotal innovation of ResNet-50 lies in the incorporation of skip connections, or residual connections, facilitating the direct flow of information from earlier to later layers. This addresses the vanishing gradient problem, enabling the training of deeper networks with heightened accuracy. Additionally, ResNet-50 employs bottleneck blocks, efficiently curbing computational costs while preserving robust representational power. The architecture's prowess is evident in its state-of-the-art achievements across various benchmarks, including the renowned *ImageNet* dataset (http://www.image-net.org), underscoring its effectiveness in image classification tasks.

Once the three models of ResNet50 are trained, it is used to predict age, gender, and emotional labels for the developed child trafficking dataset. The outputs of these three models in terms of age, gender, and emotion are stored in a comma-separated value (*.csv*) file as shown in Figure 2. Finally, to detect whether the picture denotes girl child trafficking or not, we developed another model using an SVM binary classifier as shown in Figure 2. SVM, a supervised learning approach, stands out for its proficiency in handling classification and regression tasks, especially adept at tackling intricate problems involving high-dimensional data and non-linear decision boundaries. The primary goal of SVM is to identify an optimal hyperplane that maximizes the separation between distinct classes within the input space. The determination of this hyperplane involves the selection of support vectors, representing data points closest to the decision boundary. The mathematical representation of any hyperplane can be expressed through the following equation -

$$W^T X + b = 0 \tag{1}$$

In this equation, W^T represents the transposed weight vector, X signifies the dimension of the data, and b denotes the bias term. SVM utilizes a kernel function to transform input data into a higher-dimensional feature space, facilitating non-linear classification. Moreover, SVM ensures resilience against overfitting by regulating the balance between model complexity and classification error through a regularization parameter.

We train the SVM model using the labeled data from the .csv file. Subsequently, we tested the performance of the SVM model on the same dataset and evaluated its ability to accurately classify images of girl child trafficking. The proposed approach combines the strengths of the age, gender, and emotion classification ResNet50 model with the SVM model for the analysis of .csv file data and provides a comprehensive solution for understanding and predicting the attributes of age, gender, and emotion in girl child trafficking images. Experimental results observed in our study are discussed as follows.

For our research, we compiled a dataset consisting of labeled images to train and evaluate the girl child trafficking detection model. We meticulously collected images depicting child trafficking cases, which were assigned a label of 1, representing positive instances. Additionally, we gathered images of normal child subjects to serve as negative instances and assigned the label 0. The carefully curated dataset with distinct labels allowed us to train our machine learning model to distinguish between trafficking-related images and non-trafficking images effectively.

4. Experimental Results

All the proposed models were developed using *Python 3.9.11 in NVIDIA GPU, CUDA version 11.2*. In Table I, we reported the experimental results of the three models – age, gender, and emotion. It may be noted that these models were developed using ResNet50. All these models were run up to 50 epochs. Out of the whole dataset, as reported in Table I 75% of the data was used for training, and the rest 25% was used for validation purposes. In Figure 3, we presented the training and testing accuracy observed by the three different models - age, gender, and emotion. Summarized results of observation are reported in Table I. It may be noted that the three models perform quite efficiently with a training accuracy of 89.38%, 93.51%, and 85.75%, respectively, while the testing accuracies were 80.23%, 76.298%, and 85.73%, respectively. Following these three models, the experimental results of the proposed final model developed using a Support Vector Machine to detect automatic girl child trafficking are discussed as follows.

We planned to train the final model of child trafficking developed by SVM. The one thousand four hundred fifty-six child trafficking images are considered for this purpose. It may be noted that the images are labeled by us (trafficked image = 1 and non-trafficked image =0). For all these images, we computed the age, gender, and emotions using the three developed models discussed earlier. Following this, the SVM is used to determine whether the image is of a girl child trafficking or not. The three models of gender, age, and emotion were applied to those images, and their gender, age, and emotions were predicted and stored in a .csv file. Out of one thousand four hundred fifty-six labeled

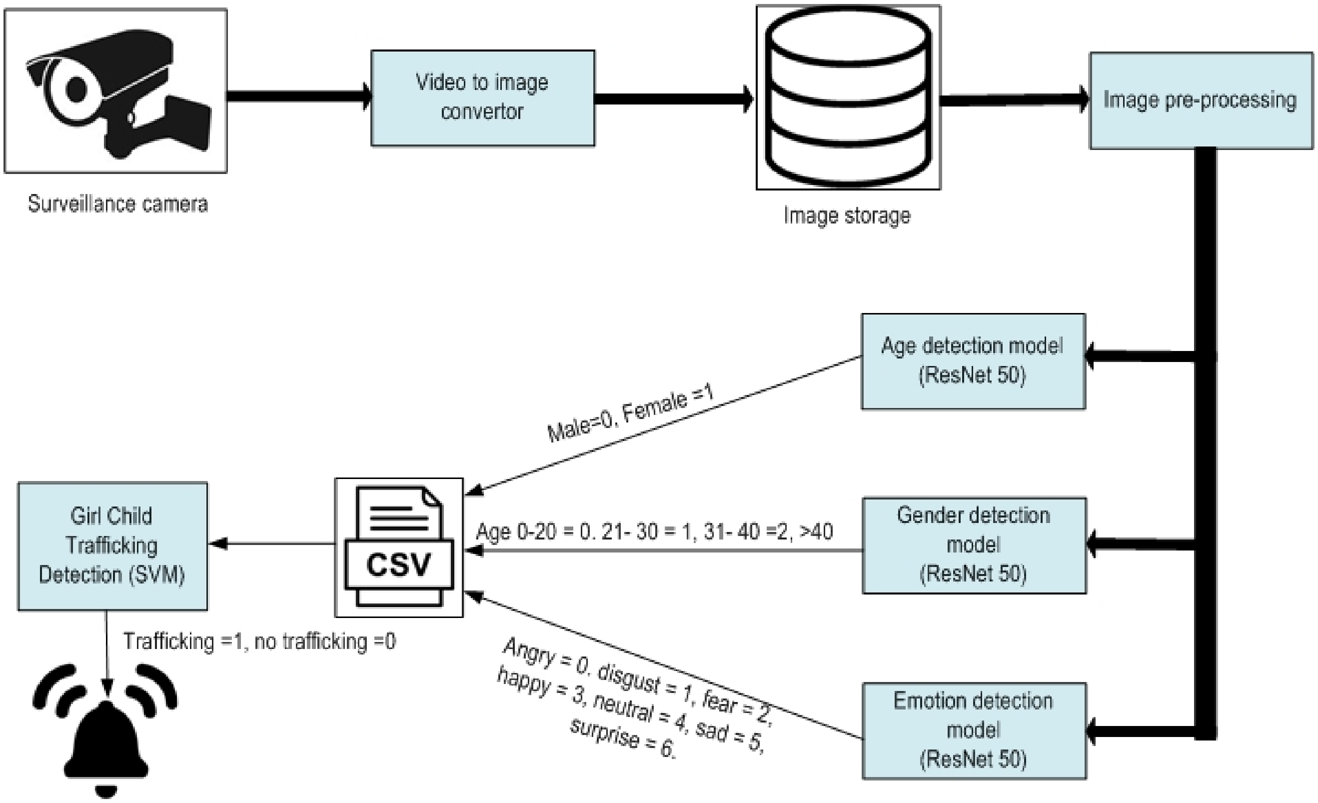


Figure 2. System architecture of our proposed model.

TABLE I. Experimental results of gender, age, and emotion detection model.

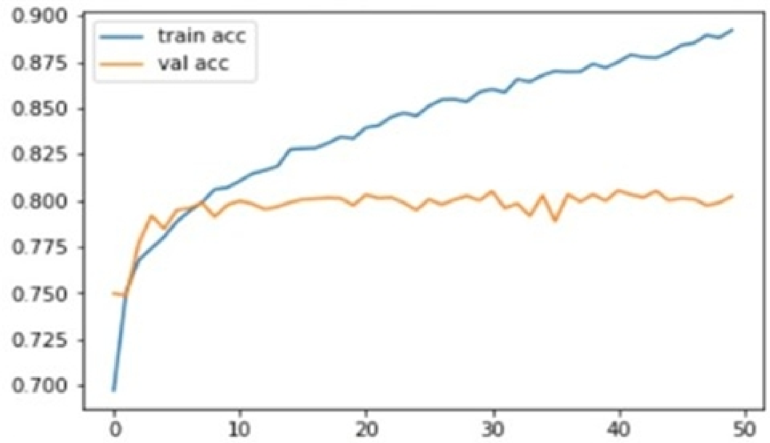
Model	No of Samples	Train Accuracy	Test Accuracy
Gender Detection	27000	89.38%	80.23%
Age Detection	9000	93.51%	76.298%
Emotion Detection	35000	85.75%	85.73%

images 75% (thousand ninety-two) are used for training the SVM, and the rest 25% (three hundred sixty-four) are used for testing purposes. The dataset developed by us is used for this purpose. We adopted a K-fold cross-validation technique (K= 6) in SVM to avoid over-fitting. In Table II, we reported the results of it. It may be noted that the proposed model works efficiently with a test accuracy of 93.13%. A Cohen Kappa score of 0.81 also illustrates the suitability of the proposed model. The confusion matrix of the obtained results is shown in Figure 4.

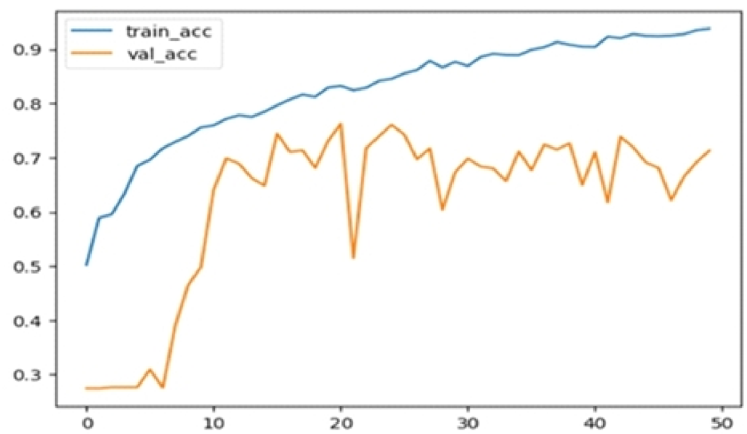
In our research, we have used various machine learning algorithms, including closest K neighbors (KNN), logistic regression, negative bays, random forests, and support vector machine (SVM), to identify the most appropriate approaches to the girls' child trafficking detection model. After careful evaluation, SVM has emerged as the most promising algorithm, showing high precision in classifying potential trafficking indicators from images. Its ability to handle complex decision boundaries and nonlinear relationships within data makes it suitable for this task. As a result, we have chosen SVM as the main classification in the model and have greatly contributed to its overall efficiency and performance. SVM's proficiency in identifying optimal decision boundaries, accommodating nonlinear data, and effectively navigating high-dimensional spaces has rendered it a favored choice across diverse classification tasks. This encompasses applications such as detecting girl child trafficking from images. The pivotal factors influencing the selection of SVM as the core algorithm in the proposed model were its exceptional accuracy and robust generalization capabilities.

5. Discussion

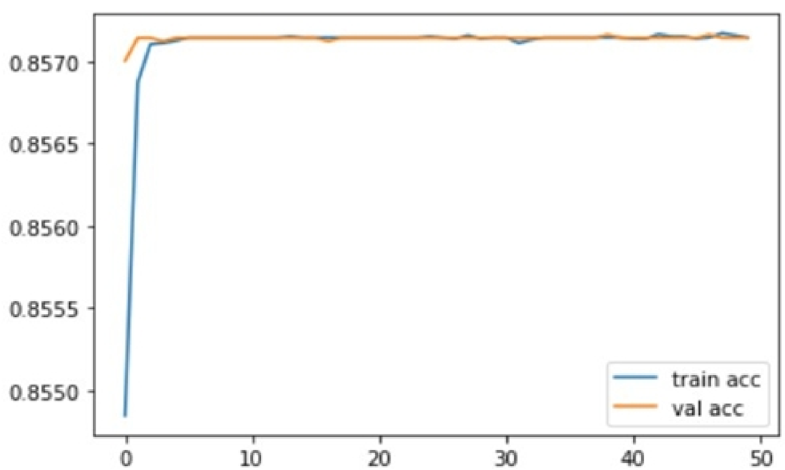
In this work, the importance of certain key factors, namely age, gender, and emotion, in predicting instances of child trafficking is used. By leveraging machine learning techniques, particularly the successful implementation of a Support Vector Machine (SVM) model, we developed



(a)



(b)



(c)

Figure 3. Train and test accuracy - a) gender, b) age, and c) emotion.

TABLE II. Results of girl child trafficking.

	%Accuracy	Precision	Recall	F1-Score	K-fold Score	Cohen Score
K-Nearest Neighbour	92.58%	0.9328	0.9258	0.9255	91.01%	0.8518
Logistic-Regression	91.48%	0.9194	0.9148	0.9146	91.19%	0.8298
Naive-Bayes	91.75%	0.9217	0.9175	0.9174	89.63%	0.8353
Random-Forest	91.75%	0.9217	0.9174	0.9174	93.76%	0.8353
SVM	93.13%	0.9107	0.9093	0.9092	88.90%	0.8185

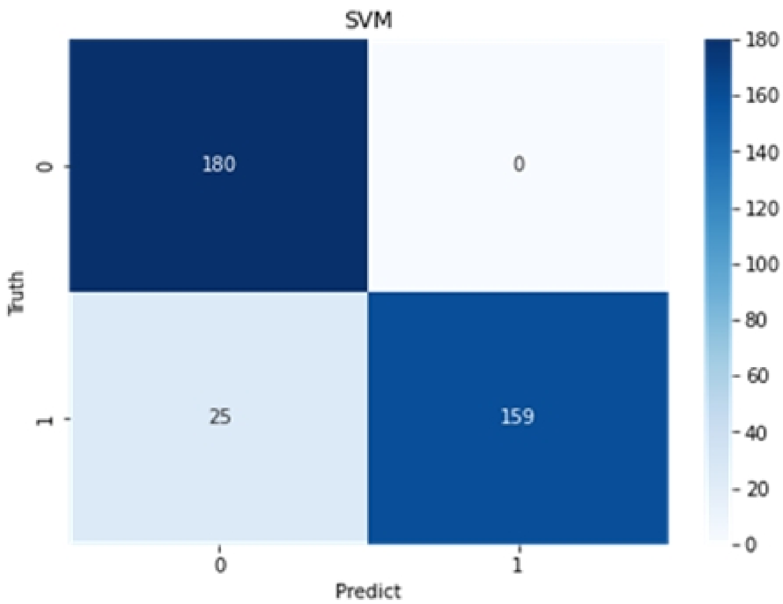


Figure 4. Confusion matrix of girl child trafficking using SVM.

our proposed model. Understanding and incorporating three factors - age, gender, and emotion into predictive models can contribute to the identification and prevention of child trafficking.

The application of a Support Vector Machine model exemplifies the practical use of machine learning in this particular context. SVM is a potent algorithm that can classify and predict results based on labeled training data. Its effectiveness in this scenario indicates that machine learning can play a crucial role in developing efficient tools to combat child trafficking.

Moreover, there is a scope for refinement and expansion of these models and datasets. This implies a commitment to continuous improvement and adaptation to enhance the accuracy and efficiency of predictive models. The dynamic nature of child trafficking requires constant updates and improvements to stay ahead of evolving patterns and tactics employed by traffickers.

The statement emphasizes the global importance of the issue and highlights that the efforts made to address child trafficking are part of a broader initiative aimed at safeguarding the well-being and rights of children worldwide. By utilizing data-driven approaches, the statement suggests that technology and analytics can be employed to make significant progress in combating child trafficking. The ultimate objective is to eliminate this heinous crime and create a safer and brighter future for all children.

During the development of our work, we noted a few limitations of our work. The utilization of age, gender, and emotion as features in the SVM model enables the identification and classification of potential instances of child trafficking. We acknowledged the factors such as the quality and representativeness of the dataset as potential constraints. Our findings have far-reaching implications for child protection organizations, law enforcement agencies, and policymakers, providing them with valuable insights to strengthen their strategies and interventions. By continuously improving these models and data sets, we can enhance the accuracy and effectiveness of predictions, empowering authorities to take proactive measures in detecting and preventing child trafficking. It is important to recognize the limitations of this research, such as the quality and representatives of the dataset. While age, gender, and emotion are significant factors, there may be other variables that could contribute to a more comprehensive understanding of child trafficking prevention.

Several strategies have been implemented to reduce the likelihood that the system will mistakenly identify a person as a victim of trafficking. First, we used a dataset containing diverse representative examples of non-trafficking cases to refine the machine learning model. This improved the model's ability to distinguish subtle differences between real trafficking indicators and non-trafficking elements. Secondly, we have introduced a confidence threshold that allows the model to classify instances only when the prediction is highly sure. The low-confidence predictions should be flagged for human review and human supervision in the decision-making process. By combining these techniques, we sought to minimize false alarms and improve the overall accuracy and reliability of the system for detecting trafficking in girls.

The proposed model greatly relies on the detection of age, gender, and emotions. In a real-time scenario, the low quality of the video may lead to poor image extraction and may harm the performance of the proposed model. Extracting facial expressions from an image captured from highly populated areas - bus terminals, and train stations may be difficult due to the occlusion, and different angles of cameras. Future research in this direction may also be interesting to explore for prospective researchers. The proposed model of girl child trafficking may be considered as a generalized version of trafficking by ignoring the gender model. Even by considering the boys' gender, the model may be treated as a boy child trafficking. Future research may be carried out to explore the potentiality of our work.

6. Conclusion and Future Works

In summary, this research underscores the potential of age, gender, and emotion as pivotal features in predicting child trafficking. The successful integration of an SVM model demonstrates the practical application of machine learning in addressing this urgent problem. Through the continuous advancement of these models and datasets, we can make substantial progress in the fight against child trafficking, working towards a world where every child is safe, protected, and free from exploitation.

This research proposes a comprehensive solution to analyze and predict the age, gender, and emotional attributes of images related to girl child trafficking. The models used in this study use the ResNet-50 architecture for age, gender, and emotion classification, and the SVM model to detect child trafficking of girls.ResNet-50 is a powerful convolutional neural network known for its success in computer vision tasks. Its main innovation, skips connections, and allows the flow of information between earlier and later layers, alleviating the problem of disappearing gradients and enabling effective training of deep networks. The use of bottleneck blocks further reduces computational costs while maintaining strong representational power. By training ResNet-50 on separate datasets for age, gender, and emotion, it can accurately predict these attributes. To detect girl trafficking, a binary SVM classifier is used.SVM is suitable for complex and high-dimensional data with nonlinear decision limits. It finds the optimal hyperplane that separates the different classes at the maximum, and the kernel function enables non-linear classification. The regularization parameter of SVM ensures robustness against overfitting. The research combines the outputs of the ResNet-50 models and feeds them into the SVM classifier to accurately classify images as representing or not trafficking of girls. The strengths of the proposed approach lie in the ability to extract age, gender, and emotional information from the images, which can be crucial to identifying potential victims and getting an understanding of the nature of the problem. In addition, the SVM classifier adds another layer of analysis to distinguish patterns associated with trafficking. Overall, this research provides a promising framework for dealing with the key problem of girl trafficking using the latest deep learning and machine learning techniques. Our aim is to implement the proposed model in real-time scenarios, focusing on applications related to closed-circuit television (CCTV) and video analysis. We intend to expand our data set to incorporate a wider range of scenarios and improve the reliability of the model. Furthermore, we aim to explore the feasibility of applying models in dynamic modes so that they can adapt and learn from the evolving environment. To achieve real-time settings, continuous data collection and refinement of the model are required. Through these approaches, we aim to improve the applicability and performance of models in practical scenarios. A machine-learning girl-child trafficking model can find other uses beyond surveillance cameras. It can be integrated into social media platforms to automatically detect and recognize potential trafficking incidents in images and videos shared online. Non-governmental organizations can use this model to identify potential victims and gather evidence for investigations. The border and immigration authorities can use the model at checkpoints to screen travelers' images for possible trafficking indicators. The advertising platform can use the model to monitor and block advertising that promotes trafficking or child exploitation. Educational institutions can ensure a safe environment by scanning images shared on their platform. Mobile applications for the protection of children can integrate the model for content analysis in real-time. Criminal justice agencies can use the tool as a screening tool to process a large number of images and identify potential suspects in investigations. This model helps to identify and protect vulnerable children during humanitarian missions. In addition, its insights can be used in public awareness campaigns to raise awareness of the trafficking of girls.

References

[1] B. Abirami, T. Subashini, and V. Mahavaishnavi, "Gender and age prediction from real time facial images using cnn," *Materials Today: Proceedings*, vol. 33, pp. 4708–4712, 2020.

[2] G. Antipov, M. Baccouche, S.-A. Berrani, and J.-L. Dugelay,

"Effective training of convolutional neural networks for face-based gender and age prediction," *Pattern Recognition*, vol. 72, pp. 15–26, 2017.

[3] A. S. Al-Shannaq and L. A. Elrefaei, "Comprehensive analysis of the literature for age estimation from facial images," *IEEE Access*, vol. 7, pp. 93 229–93 249, 2019.

[4] T. V. Janahiraman and P. Subramaniam, "Gender classification based on asian faces using deep learning," in *2019 IEEE 9th International Conference on System Engineering and Technology (ICSET)*. IEEE, 2019, pp. 84–89.

[5] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 34–42.

[6] C. H. Nga, K.-T. Nguyen, N. C. Tran, and J.-C. Wang, "Transfer learning for gender and age prediction," in *2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan)*. IEEE, 2020, pp. 1–2.

[7] E. W. Teh and G. Taylor, "Apparent age estimation with relational networks," in *2019 16th Conference on Computer and Robot Vision (CRV)*. IEEE, 2019, pp. 89–96.

[8] S. A. Khan, M. Ahmad, M. Nazir, and N. Riaz, "A comparative analysis of gender classification techniques," *International Journal of Bio-Science and Bio-Technology*, vol. 5, no. 4, pp. 223–244, 2013.

[9] A. Dhomne, R. Kumar, and V. Bhan, "Gender recognition through face using deep learning," *Procedia computer science*, vol. 132, pp. 2–10, 2018.

[10] S. Gupta, "Gender detection using machine learning techniques and delaunay triangulation," *International Journal of Computer Applications*, vol. 124, no. 6, 2015.

[11] D. Duncan, G. Shine, and C. English, "Facial emotion recognition in real time," *Computer Science*, pp. 1–7, 2016.

[12] A. Jaiswal, A. K. Raju, and S. Deb, "Facial emotion detection using deep learning," in *2020 international conference for emerging technology (INCET)*. IEEE, 2020, pp. 1–5.

[13] W. Zheng, H. Tang, Z. Lin, and T. S. Huang, "Emotion recognition from arbitrary view facial images," in *Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part VI 11*. Springer, 2010, pp. 490–503.

[14] P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller, and S. Zafeiriou, "End-to-end multimodal emotion recognition using deep neural networks," *IEEE Journal of selected topics in signal processing*, vol. 11, no. 8, pp. 1301–1309, 2017.

[15] A. I. Siam, N. F. Soliman, A. D. Algarni, A. El-Samie, E. Fathi, A. Sedik *et al.*, "Deploying machine learning techniques for human emotion detection," *Computational intelligence and neuroscience*, vol. 2022, 2022.

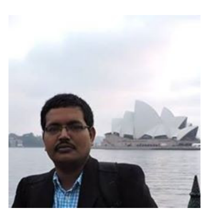
[16] I. Maglogiannis, D. Vouyioukas, and C. Aggelopoulos, "Face detection and recognition of natural human emotion using markov random fields," *Personal and Ubiquitous Computing*, vol. 13, pp. 95–101, 2009.

[17] H.-W. Ng, V. D. Nguyen, V. Vonikakis, and S. Winkler, "Deep learning for emotion recognition on small datasets using transfer learning," in *Proceedings of the 2015 ACM on international conference on multimodal interaction*, 2015, pp. 443–449.

Dhrubajit Kakati Dhrubajit Kakati is currently pursuing a B.Tech degree from the Central Institute of Technology Kokrajhar, BTR, Assam, India. His research interests include Machine Learning, Deep Learning, and Image Processing.

Rajarshi Seal Rajarshi Seal is pursuing a B.Tech degree from the Central Institute of Technology Kokrajhar, BTR, Assam, India. His research interests include Machine Learning, Deep Learning, and Image Processing.

Tania Sarkar Tania Sarkar works as an Assistant Professor in the Department of Computer Science and Engineering at the Central Institute of Technology Kokrajhar, BTR, Assam, India. Her research interests include Machine Learning, Deep Learning, IoT, and Image Processing.



Dr. Ranjan Maity Dr. Ranjan Maity has a B.Tech degree from Kalyani University, an MS from IIT Kharagpur, and a PhD from IIT Guwahati. He currently works as an Assistant Professor in the Department of Computer Science and Engineering at the Central Institute of Technology Kokrajhar, BTR, Assam, India. His research areas include usability of interfaces, machine and deep learning, and AR/VR. Dr Maity is a Senior Member of both IEEE and ACM.