International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Identification of Age and Gender through Crowd Analysis

Akshata A Kokate[1], Shivani S Pawar[2], Mr. Abhishek Nazare[3]

[1]MCA Student, [2]Associate Professor

Department of MCA, K. L. S. Gogte Institute of Technology, Belagavi, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India

ABSTRACT:

These days combining age and gender prediction with facial recognition has greatly improved crowd surveillance making it a vital yet challenging task this work presents a system that simultaneously identifies individuals estimates their age and predicts their gender using well-known models like face net resent support vector machine age net and gender net we tested this method on a challenging new data set with many faces unlike existing facial recognition technology that focuses only on static data sets of known faces our system can recognize new faces and update its database for continuous monitoring making it more suitable for ongoing crowd surveillance. as fresh faces are identified in our suggested method the system saves them with a special label and updates on a regular basis so that it can identify them in scans to come every time a new face is identified it is inefficient to extract facial features from the complete data set we suggest a technique for incremental feature extraction to lessen the computational burden in order to resolve this. our system successfully identified per-trained identities with 49 accuracy assessed age with 665 accuracy and predicted gender with 9354 accuracy when evaluated on the suggested data set we discovered that these per-trained models can be sensitive and unreliable in unpredictable situations such as abrupt changes in lighting.

KEY WORDS: Age estimation, crowd monitoring, deep learning, facial recognition, gender prediction

1. INTRODUCTION:

Facial recognition software has become increasingly prevalent in artificial intelligence as it is necessary to detect and validate peoples biometrics using learning algorithms now identity theft security in general and other factors are significantly increasing the demand for facial recognition face identification tasks typically focus on recognizing pre-trained identities or those that are already familiar and rarely on identifying new appearances where there are no familiar identities this becomes a major bottleneck in real-time and large scale public surveillance that the same-to-same is always in the crowd scene face recognition age estimation face transformation draw your resume rightful close to you in some reasons gender prediction models at state-of-the-art accuracy however most these models were not tested in practical reality for example different faces in a photo unbalanced datum this in this context this paper presents a deep learning-oriented approach system to the facial recognition age detection and adult content detection all performed at the samedefvarining gender prediction the deep learning-based method proposed in this paper aims to concurrently execute age estimation face recognition and gender forecasting this approach makes use of the currently deep learning and machine learning models that are currently accessible in place of own-created models the proposed pipeline is made up of support vector machines facenet[21] for obtaining facial features and resnet-based detector [20] for face detection age estimate using agenet [22] identity categorization and predicting gender with gendernet to reduce the computational load we also develop an incremental feature extraction method rather than extracting the characteristics of every face in the dataset the suggested method saves identities for pre-processing such as face alignment clustering and labelling whenever new identities are found in the frame following pre-processing the new faces will have their features that is 128-d embeddings using facenetextracted it wont be necessary to extract all of the embeddings again because the recently extracted embeddings will simply be combined with the preexisting old embeddings. the combined embeddings will then be used to train an svm model the model can be updated on a regular basis with this strategy lastly in order to assess the suggested solution we offer a genuine demanding multi-face test set with photos captured in various lighting scenarios and stances the remainder of the paper is structured as follows the connected works are examined in section after that section discusses the research technique the results are examined and discussed in section section finally brings the paper to a close

2. RELATED TO WORK:

Works by are regarded as pioneers in the field of automatic facial recognition they recommended employing specialised edge and contour detectors to locate a set of facial landmarks and measure the distances and relative locations between them we call these geometry-based techniques subsequently the statistical sub spaces methodologies known as principal component analysis pca and linear discriminant analysis lda became more famous holistic approaches is another name for these in the concept of using pca on a set of training face photos to identify eigen vectors also referred to as eigen faces was initially offered. the distribution of data is most variable when eigen vectors are used price et al suggested an all-encompassing approach based on lda support vector machines svms have also been used to recognise faces in a method connected to pca and lda called locality preserving projections lpp was presented the suggested joint bayesian technique [23] proposes to describe a facial picture instead of image differences by summing two independent gaussian variables this approach on the labelled faces in the wild lfw dataset yielded the best accuracy of any holistic method to yet at according to a proposal in feature-based methods compare these local properties in different face photos the term modular eigen faces method applied to this the ebgm approach which is also widely used was introduced in it utilised a graph of nodes containing gabor wavelet coefficients extracted around a predetermined set of facial landmarks to depict a face presented an approach that outperforms the ebgm by substituting histograms of oriented gradients hog for gabor wavelet features various feature-based approaches have focused on learning local features from training samples combining feature-based and holistic approaches led to the creation of the hybrid approach a popular approach involves extracting local features such as lbp and scale-invariant feature transform sift and projecting them onto a lower-dimensional and discriminative subspace such as pca or lda other approaches simply combine the two techniques without any interaction thanks to improvements in computer processing power and data storage capacity deep learning dl based facial recognition techniques have gained popularity recently convolutional neural networks [22] [31] cnns which are end-to-end trainable models having the benefit of being trained [25] with a huge quantity of data and are able to learn the variations of face representations existent in the training data are the most popular form of dl approaches for fr after being trained each subject belongs to a class that can identify faces that arent in the training dataset [24] according to one of the cnn [25] models training methodologies another popular method for learning bottleneck features is to optimise the distance metric between triplets of faces large-scale data training made dl facial recognition techniques state-of-the-art on the lfw benchmark facebooks deepface achieved an accuracy of 97.35% the cnn with softmax loss [26] was trained using an enormous dataset of 4.4 million faces from 4030 participants trained 60 distinct cnns on patches containing 202,599 face photos of 10,177 celebrities and the findings were identical to those of [1] conducted an extensive analysis of several cnn architectures and the findings indicated that a 100-layer resnet produced the best trade-off between accuracy speed and model size. 200 million face identities and 800 million picture face pairs were used by google to train the facenet [2] cnn model their method involved applying a triplet based loss at various layers comparing a third distinct face z and a pair of identical faces x y in order to move x closer to y than z at 9963 this methods accuracy on lfw is among the highest available right now the most widely used approach for resolving facial recognition issues at the moment is facenet which is employed by many academics in this field [15]. Issues related to automatically determining an object's age have been discussed for a while. An early method of age nose, eyes, and face. The ratio between them is computed by estimating distances and sizes in order to estimate age using traditional techniques following localization [27]. Age determination by facial geometry analysis, such as the pipeline [28] employed in , has been a popular practice since the 1990s research. Partial Least Square (PLS), Canonical Correlation Analysis (CCA), and Biologically Inspired Features (BIF) were merged, for instance, in face photos were previously included in bif which opened the door for following studies like which showed that automatic approach could match human performance the foundation of techniques prior to cnns was two-stage pipeline-like feature extraction like lbp [30] followed by classification using an svm or multilayer perceptron mlp conversely cnns carry out the previously indicated procedure in a single step learning both the extraction and the categorization of age groups or by executing age regression

In contrast to age analysis gender recognition using neural networks such as the method of was first suggested in the early 1990s two neural networks were proposed in an autoencoder and a classifier whose input was the auto encoders encoded output layer the largest flaw in it was that it relied on hand cutting scaling and facial rotation in the photo taken under supervision inspired by age estimation approach pipelines [32] based on a feature extractor and stacked classifier were presented in gender recognition was achieved in by fine-tuning a [33] pre-trained network and then training an svm using deep features computed by cnn the same cnn-based techniques used in [22] to determine gender were also applied demonstrating [29] [33] that cnns are capable of performing tasks by only altering the data utilised for learning and nothing else researchers have conducted a number of experiments on age and gender recognition using cnn-based techniques in recent years and many of the works [22] produced state-of-the-art outcomes and unlike gender and age prediction it is clear from the aforementioned studies that facial recognition is done independently in this study we combine the contributions of to offer a unified system for simultaneous face detection identification age estimation gender prediction and prediction since age and gender in addition to identity can be used to classify peoples behaviour and monitor the crowd more effectively. to effectively train an svm model we also provide an incremental embeddings extraction technique this researchs main goal is to assess the suggested system utilising our recently developed demanding and realistic test dataset the test datasets photographs feature a range of people in different poses and lighting circumstances for example three people in one picture.

3. METHODOLOGY

For a variety of purposes including facial recognition deep learning has emerged as a key tool for detection and classification identification of the individual age estimation and gender prediction are the main topics of this essay one can determine age from an image or a real-time stream using age estimation which is either a classification problem or a regression problem in contrast facial recognition and gender prediction involve classification issues here we utilise the face detector based on resnet the person identification method facenet in conjunction with support vector machine the age estimate method agenet [22] and the gender prediction method gendernet [22] the next subsections provide brief descriptions of facenet agenet and gendernet the operations flow is then explained divided into two sections training and inference the training process is detailed in the first section followed by a description of the inference process in the second section.

FaceNet

Google researchers developed and released facenet in 2015 to tackle challenges related to face detection and verification facenet utilizes a deep convolutional network to directly optimize embeddings this differs from earlier deep learning methods that used an intermediate bottleneck layer the facenet technique employs a one-shot learning method converting face images into 128-dimensional vectors in euclidean space also known as embeddings. random and dissimilar photos are located widely away in compared to related images which allows for face detection and verification utilising facenet embeddings as feature vectors the deep convolutional neural network normalisation which creates the face embeddings after the deep convolutional network and the triplet loss input further improve the process are both included in the facenet network architecture images consist of closely cropped facial pictures. a compact 128-dimensional embedding is trained from facenets output using a triplet-based loss function three imagesan anchor a positive and a negativeare required to compute triplet loss of these photos the positive image and the anchor are the same but the negative image and the anchor are not this technique makes sure the network learns to generate embeddings in the embedding space where dissimilar images are separated from each other and comparable images are closer together.

3.2 AgeNet and GenderNet

Levi and hassner developed the agenet and gendernet model agenet and gendernet use a basic design similar to alexnet eight age groups 0-2 4-6 8-13 15-20 25-32 38-43 48-53 60 and two gender brackets male female are learned by this architecture there is variability in these age groups this is a result of the model being trained using the adience dataset [42] the age ranges are precisely as defined by these brackets in the adience dataset. accurately estimating an individuals age regression problem is never easy due to factors including heredity physical attributes makeup and the aftermath of plastic surgery because of this agenets age estimation algorithm detects age groups for categorization which somewhat simplifies the work as a result of its binary classification gender prediction is however relatively.

Simpler there are just three convolutional layers and two fully linked layers with a limited number of neurons in the agenet-gendernet model design. a corrected linear operation and a pooling layer come after each of these three convolutional layers using local response normalisation the first two convolutional layers likewise adhere to normalisation two completely connected layers with 512 neurons per are added at the end with the adience dataset this approach has one of the highest accuracy to date gender prediction obtained 868 exact accuracy and age estimation achieved 847 one-off accuracy both of which greatly beat the then state-of-the-art approaches.

3.3 Flow of Training Operations

Because the proposed system must identify new identities in real-time during inference and add them to the constantly growing dataset the dataset is essential for training a machine learning or deep learning model this allows the system to recognise these identities in subsequent inference runs for our recognition assignment the manual dataset creation method is unfeasible the automated dataset creator system or adcs is our recommended answer to this problem. the face image of each identity that is identified during an inference run is saved by adcs which then clusters and aligns the faces to assign each person a unique id periodically the svm model will undergo retraining and the 128-d embeddings of the newly added faces will be extracted based on the updated dataset consequently in the event that those new identities are discovered once more at a later time step the system will be capable of identifying them and assigning the appropriate unique id. We already have a pretrained model that we can use for age estimation and gender prediction, so we don't need to train it. In conclusion, Figure 1 illustrates the seven phases that make up the ADCS. Here is a list of these:

1) Find faces that are unfamiliar during inference, then mark them for additional analysis.

2) Orient every face in queue. Assign each face a unique ID and group them together.

3) Then include them in the dataset. Remove the new faces embeddings.

4) Combining the new and old embeddings together is the first step.

5) Retrain SVM model.

However the system needs to be trained or initialised using an initial dataset before the inference operation can be performed eleven classes of known identities and one class of unknown identities make up our initial dataset randomly gathered frontal face photos of persons are included in the uncertain identity class conversely the remaining 11 classes of recognised identities feature pictures of 11 distinct individuals that were captured in various lighting scenarios and with diverse body positions a representative of the 11 recognised classes is shown in figure 2 the adcs workflow is described in depth in the following subsections

(ADCS)

3.4 Detecting and Saving Novel Identities

First we identify new identities in streaming video the face area of interest roi is extracted from the video stream by feedforwarding it to the resnet face detection model during inference afterwards the 128-dimensional face embeddings are produced by feedforwarding the face roi to the cnn model for classification the svm model receives these facial embeddings as input the identity is then subject to a conditional criterion whereby it is designated as unknown if it differs more from the trained known classes than it does from them if not it receives a label from one of the pre-trained known classes that currently exist.

3.5 Face Aligment

It is necessary to align the newly discovered identities in order to improve the clustering outcomes identifying the geometric facial structure in digital images and attempting to maintain canonical facial symmetry based on translation size and rotation is known as face alignment to convert the image into an output coordinate system if there are multiple facial landmarks input coordinates is the primary goal the objective of this method is to ensure that faces remain visually identical for better performance with the facial area positioned in the centre of the image and the eye in a horizontal position. in order to identify the left and right eye areas a face landmark model is first used finding each eyes centre which can be a parameter to adjust the rotation is the second stage it then computes the angle of face rotation using the left eyes x coordinate the fourth step determines the needed or desired right eye step 5 calculates the eye-center or midpoint between two eyes utilising the rotation matrixwhich is produced by combining all of the previously listed parametersthe last stage is aligning the face.

3.6 Face Clustering

Following face alignment the faces must be clustered in order to mark each group of related faces with a distinct id face clustering uses unsupervised learning which comprises solely faces without classes whereas face recognition uses supervised learning for categorization we extract discriminative representation for the faces in the clustering job which is a crucial requirement for the clustering technique this is taking each image that will be used as a face representation and extracting a 128-dimensional feature vector called encoding. a streamlined version of resnet dnn deep neural network is used for this procedure density based spatial clustering of applications with noise dbscan is the method utilised for the clustering dbscan gathers neighbouring points that are packed together from n-dimensional space as a result a single cluster with nearby points is produced additionally dbscan handles outliers well dbscan is used to cluster the face encodings into distinct clusters once they have been extracted outliers are eliminated here every distinct cluster denotes a distinct class label with a single label assigned to every face within a cluster upon completion of clustering the faces are prepared for feature extraction figure 4 shows how two sample clusters performed

3.7 Incremental Facial Features Extraction and SVM Training

The full dataset is usually processed using facenet in order to obtain face embeddings for each identity it is now possible to detect and localise faces in the photographs by putting them into the resnet model which creates the face area of interest roi following receipt of the face roi the facenet model creates a 128-dimensional facial embedding vector for every face following that these embeddings and the labels that correlate to them are used to train an svm model but using facenet to create embeddings for every face that is found is not feasible because of how time-consuming and computationally intensive this procedure is we provide an incremental feature extraction technique to address this issue this method uses the old embedding data and the new embedding data to train an svm exclusively extracting the embeddings of newly discovered faces consequently rather than creating the embeddings for each identity from scratch we may just append the new embeddings to the existing embeddings the model can be applied in practical settings thanks to the suggested strategy which lowers computing load and conserves memory the process is shown in figure 5

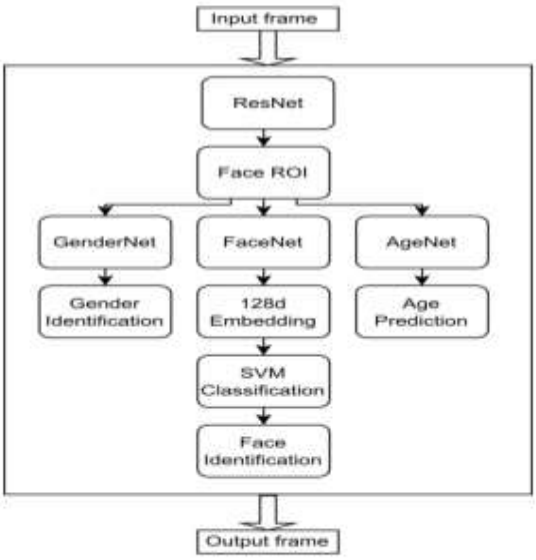

Figure 3 Flow of inference operation

4. RESULTS AND DISSCUSSION

In order to assess the suggested approach we put forth a realistic and demanding test dataset that includes photos captured under both bright and low lighting conditions a real-world video feed may have many faces in each frame we have added pictures to our test dataset that range in the number of facesfrom one to fourin order to represent this.

4.1 Identity Recognition Results

4.1.1 Accuracy

Figure 8a shows that the accuracy for a single individual in low light is 56 when there are several persons in the frame the accuracy decreases in high light one person may obtain 68 accuracy and as the number of persons in the frame increases accuracy rapidly decreases as seen in figure 8b when all possibilities are taken into account the average accuracy is 49%

4.1.2 Precision

Accuracy is sometimes referred to as repeatable or reliable the accuracy of identifying a single person in a group of people decreases rapidly 6344 in bright light and 5714 in low light

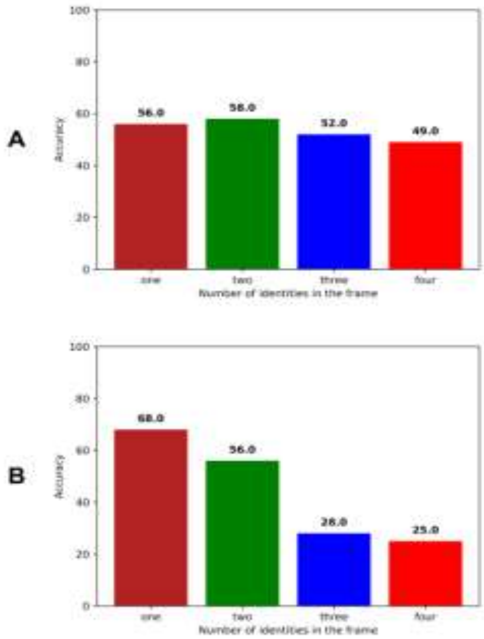

Figure 4 Accuracy in (A): dark light condition and (B): bright light condition

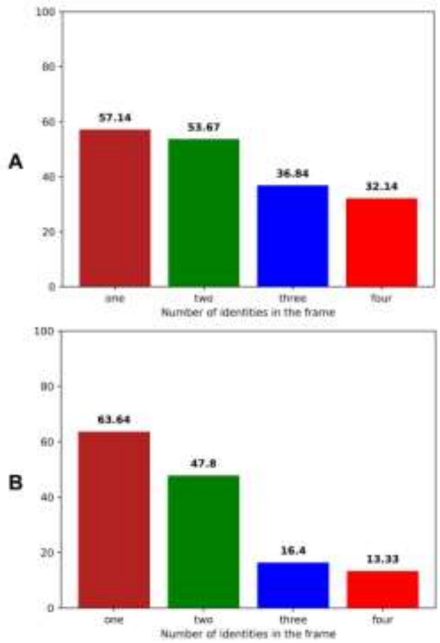

Figure 5 Precision in (A): dark light condition and (B): bright light condition

4.1.3 Recall

True positive rate is another term for recall. Figure 10 illustrates that in a dark state, the maximum recall rate is 65.20 percent, while in a bright condition, the highest value is 77.00%, which is obviously greater than in a dark condition. Similarly, 50.00% and 53.85%, respectively, are the lowest values for light and dark conditions. It follows that whereas recollection varies under both circumstances, it is more stable under low light.

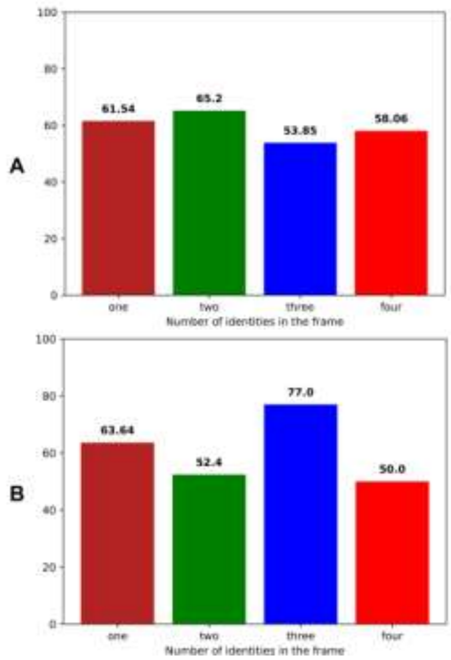

Figure 6 Recall in (A): dark light condition and (B): bright light condition

4.1.4 Specificity

Another name for true negative rate (TNR) is specificity. As shown in Figure 11A, the specificity is 50.00% for a single person in the frame in low light; as the number of people increases, it stays nearly linear and does not vary significantly. On the other hand, Figure 11B shows that in high light conditions, the specificity is 71.43% for a single person in the frame; but, as the number of people increases, it rapidly decreases before becoming linear.

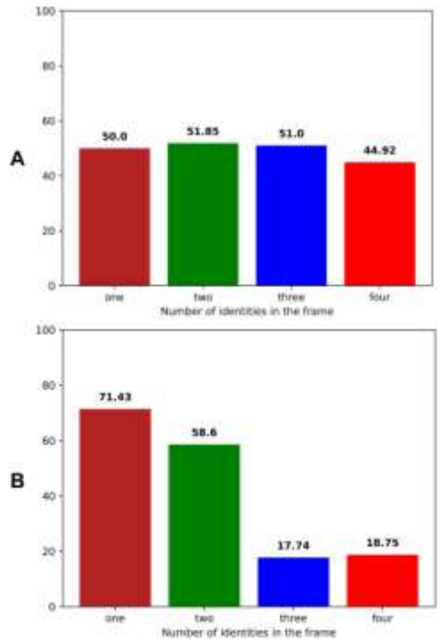

Figure 7 Specificity in (A): dark light condition and (B): bright light condition

4.2 Age estimation and Gender Prediction

Figure 12 shows that the one-off accuracy for age estimation does not follow any pattern in either of the illumination situations as can be seen the maximum one-off accuracy value is 8933 in light conditions and 7866 in dark conditions 54 is the lowest value in both conditions

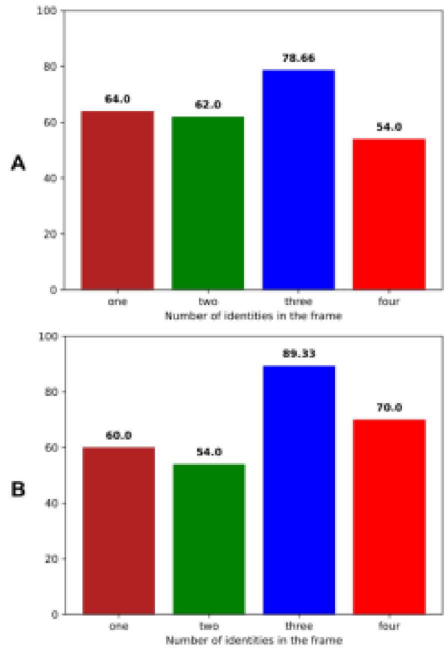

Figure 8 One-off accuracy in (A): dark light condition and (B): bright light condition

the exact accuracy for gender prediction as illustrated in figure 13 follows a linear trend in both light conditions the highest values of both dark and bright light conditions are respectively 98 and 100 the lowest value reads 81 for the dark light condition and 90 for the bright light condition

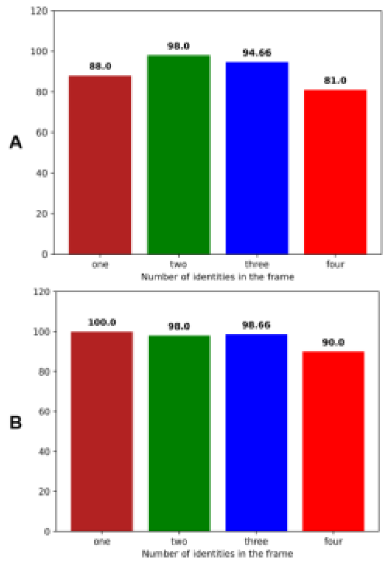

Figure 9 Exact accuracy in dark light condition (left) and bright light condition (right)

4.3 Discussion

theoretically face recognition systems should perform better in identity recognition in bright light than in dark light however rather than the general lighting condition the suggested systems identity recognition performance is mostly dependent on the lighting condition of the face area as can be seen from the test photos there was better overall lighting in bright light than in low light nonetheless compared to strong light the lighting in the face area was superior in the dark state. due to this circumstance performance suffered in bright light and improved in low light furthermore since the test dataset was not gathered in a controlled manner it is extremely difficult and highly unbalanced as most other studies often utilise a test dataset that is collected in a controlled manner the purpose of this research is to truly use a tough test set when given a clean well-controlled test dataset models typically perform quite well but when given an unbalanced uncontrolled dataset they suffer our test data for three people in a frame and four people in a frame is wildly unbalanced and has a higher proportion of unknown identities than known identities. for instance out the four people in certain pictures three are unidentified these disparities significantly impair the systems performance the overall number of known identities in 4 people in a frame and bright light condition was 20 whereas the total number of known identities was 80 this case saw 65 false positives produced by the system which quickly reduced the accuracy precision and specificity nevertheless the model produced 10 false negatives as there were substantially fewer known ids owing to this disparity recall is comparatively higher than specificity accuracy and precision. since the test dataset weve suggested is limited we think that increasing it will greatly lessen the fluctuation based on how well our suggested system performs reduced performance may also result from other variables like position variation low resolution distance from the camera and frame size additionally the study by 25 came to the conclusion that using a facenet-based system a decrease in resolution can result in a noticeable loss of accuracy also consistent with our findings is this finding. the results of the age and gender prediction indicate that the light situation and the number of people in the frame have no bearing on the one-time accuracy of age estimation or the precise accuracy of gender prediction it actually relies on the angle of the stance and the facial expression given that they are pre-trained neither model requires the addition of new training data on the adience dataset agenet obtained 847 one-off accuracy while gendernet achieved 868 precise accuracy in the initial analysis 4 following that on the utkface [34] dataset agenet and gendernet achieved 457 and 8732 precise correctness respectively in a study done by 58 63 conversely agenet and gendernet demonstrated an average one-off accuracy of 665 and 9354 respectively for age estimation on our suggested test dataset we can determine that our results are consistent with those of other studies by comparing them aligning the faces before feeding them to the models for estimate can further increase the accuracy.

5. CONCLUSION:

Our proposal in this research is to use pre-existing deep learning and machine learning models to create a system that can identify and re-identify estimate age and forecast gender of discovered individuals both known and unknown furthermore a realistic demanding multi-face test dataset has been proposed to assess the suggested approach the systems average accuracy on the suggested dataset was 49 for identity recognition 665 for age estimation and 9354 for gender prediction. it is clear from the results that the modelsfacenet and agenetare not reliable and may be vulnerable to unbalanced and erratic test datasets if certain requirements are not satisfied the system is unable to accurately identify the faces and determine the ages both the facenet and the agenet are susceptible to changes in position and facial illumination for instance in situations involving obstructed position and sudden lighting identification recognition accuracy may be as low as 25 the suggested. approach undoubtedly has to be modified in order to be used for real-world crowd monitoring. nevertheless the gendernet model outperformed the other two models with an average accuracy of 9354 which is rather high please take note that the results may differ dramatically if additional test photos are added as our proposed test dataset is somewhat tiny by using face alignment algorithms and image brightener algorithms we think the accuracy of age and identity estimation may be significantly increased to enable real-time operation on a cctv stream the suggested system also needs to be adjusted and optimised.

REFERENCES:

1. Deng, J., Guo, J., Xue, N. and Zafeiriou, S. 2019. ArcFace: Additive angular margin loss for deep face recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2019-June: 4685-4694. DOI: https://doi.org/10.1109/CVPR.2019.00482

2. Jose, E., Greeshma, M., Mithun Haridas, T. P. and Supriya, M. H. 2019, March. Face Recognition based Surveillance System Using FaceNet and MTCNN on Jetson TX2. 2019 5th International Conference on Advanced Computing and Communication Systems, ICACCS. 608-613. DOI: https://doi.org/10.1109/ICACCS.2019.8728466

3. Nyein, T. and Oo, A.N. 2019, November. University Classroom Attendance System Using FaceNet and Support Vector Machine. 2019 International Conference on Advanced Information Technologies, ICAIT. 171-176. DOI: https://doi.org/10.1109/AITC.2019.8921316

4. Golla, M.R. and Sharma, P. 2019. Performance evaluation of facenet on low resolution face images. Communications in Computer and Information Science. 839: 317-325. DOI: https://doi.org/10.1007/978-981-13-2372-0_28

5. William, I., Ignatius, D. R., Moses Setiadi, Rachmawanto, E. H., Santoso, H. A., & Sari, C. A. 2019, October. Face Recognition using FaceNet (Survey, Performance Test, and Comparison). Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019. DOI: https://doi.org/10.1109/ICIC47613.2019.8985786

6. Frearson, M. and Nguyen, K. 2020, November. Adversarial Attack on Facial Recognition using Visible Light. DOI: https://doi.org/10.36227/techrxiv.13725634.v1

7. In January 2020, Zeng, W. et al. conducted research on person recognition using FaceNet under simulated prosthetic vision. The study was published in the Journal of Physics: DOI: https://doi.org/10.1088/1742-6596/1437/1/012012

8. Cahyono, F. Wirawan, W. and Fuad Rachmadi, R. 2020, September. Face recognition system using facenet algorithm for employee presence. 4th International Conference on Vocational Education and Training, ICOVET 2020. 57-62. DOI: https://doi.org/10.1109/ICOVET50258.2020.9229888

9. Moura, A. F. S., Pereira, S. S. L., Moreira, M. W. L. and Rodrigues, J. J. P. C. 2020, January. Video Monitoring System using Facial Recognition: A Facenet-based Approach. IEEE Global Communications Conference, GLOBECOM 2020 – Proceedings. 1-6. DOI: https://doi.org/10.1109/GLOBECOM42002.2020.9348216

10. Nair, S. P., Abhinav Reddy, K., Alluri, P. K. and Lalitha, S. 2021. Face recognition and tracking for security surveillance. Journal of Intelligent and Fuzzy Systems. 41(5): 5337-5345. DOI: https://doi.org/10.3233/JIFS-189856

11. He, K., Zhang, X., Ren, S., and Sun, J. 2016. Deep residual learning for image recognition. DOI: https://doi.org/10.1109/CVPR.2016.90

12. Schroff, F., Kalenichenko, D., and Philbin, J. 2015. FaceNet: A unified embedding for face recognition and clustering. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 815-823. DOI: https://doi.org/10.1109/CVPR.2015.7298682

13. Levi, G., and Hassncer, T. 2015. Age and gender classification using convolutional neural networks. DOI: https://doi.org/10.1109/CVPRW.2015.7301352

14. Chen, d., Cao, X., Wang, L., Wen, F. & Sun, J. 2012. Bayesian face revisited: DOI: https://doi.org/10.1007/978-3-642-33712-3_41

15. Learned-Miller, E., Huang, G.B., RoyChowdhury, A., Li, H. and Hua, G. 2016. Labeled faces in the wild: A survey.DOI: https://doi.org/10.1007/978-3-319-25958-1_8

16. Taigman, Y., Yang, M., Ranzato, M. and Wolf, L. 2014. DeepFace: Closing the gap to human-level performance in face verification. Proceedings of the IEEE Computer Society DOI: https://doi.org/10.1109/CVPR.2014.220

17. Sun, Y., Wang, X. and Tang, X. 2014. Deep learning face representation from predicting 10,000 classes.DOI: https://doi.org/10.1109/CVPR.2014.244

18. Özbulak, G., Aytar, Y. and Ekenel, H. K. 2016. How transferable are CNN-based features for age and gender classification?. DOI:https://doi.org/10.1109/BIOSIG.2016.7736925

19. Guo, G. and Mu, G. 2013. Joint estimation of age, gender and ethnicity:DOI: https://doi.org/10.1109/FG.2013.6553737

20. Han, H., Otto, C., Liu, X. and Jain, A. K. 2015. Demographic estimation from face images: Human vs. machine performance. DOI: https://doi.org/10.1109/TPAMI.2014.2362759

21. Eidinger, E., Enbar, R. and Hassner, T. 2014. Age and gender estimation of unfiltered faces. DOI: https://doi.org/10.1109/TIFS.2014.2359646

22. Chen, J. C., Kumar, A., Ranjan, R., Patel, V. M., Alavi, A. & Chellappa, R. 2016. A cascaded convolutional neural network for age estimation of unconstrained faces. DOI: https://doi.org/10.1109/BTAS.2016.7791154

23. Van De Wolfshaar, J., Karaaba, M. F. and Wiering, M. A. 2015. Deep convolutional neural networks and support vector machines for gender recognition. Proceedings - 2015 DOI: https://doi.org/10.1109/SSCI.2015.37

24. Zhang, Z., Song, Y. and Qi, H. 2017, November. Age progression/regression by conditional adversarial autoencoder.DOI: https://doi.org/10.1109/CVPR.2017.463