***Department of" Bachelor of Computer Application"* Ajeenkya DY Patil University *Pune*.**

ACADEMIC YEAR 2022-2023.

SEMESTER – VI

A PROJECT REPORT ON

"AGE & GENDER DETECTION USING UTK FACE DATASET"

SUBMITTED BY

CAVIN LOBO

KRISH DHANANI

UNDER THE GUIDANCE OF

PROF. MOHIT SAXSENA

ACKNOWLEDGEMENT

The satisfaction that accompanies that the successful completion of any task would be incomplete without the mention of people whose ceaseless cooperation made it possible, whose constant guidance and encouragement crown all efforts with success.

We are grateful to our project guide **PROF. MOHIT SAXSENA.**

For the guidance, inspiration and constructive suggestions that helped us in the preparation of this project.

Lastly we are also thankful to our parents and colleagues who have helped us in successful completion of the project.

Yours Sincerely,

CAVIN LOBO

KRISH DHANANI

INDEX

PROBLEM DEFINATION:

The problem addressed in age and gender detection using the UTK Face dataset is the accurate identification of a person's age and gender based on facial images. This is an important problem in computer vision and machine learning, with applications in various fields such as surveillance, biometrics, and medical imaging.

The UTK Face dataset contains over 20,000 facial images with annotations for age and gender, making it a widely used dataset for this task. However, the task of age and gender detection from facial images is challenging due to the variations in facial expressions, lighting, pose, and other factors that can affect the accuracy of the predictions.

Therefore, the goal of this research is to develop accurate models using deep CNNs and transfer learning techniques that can predict age and gender from facial images with high accuracy. This problem is critical for improving the performance of various applications that rely on facial recognition and biometrics, such as security and surveillance systems, healthcare, and marketing.

Topic – Age & Gender Detection using UTK Face dataset.

[1]Cavin Lobo, [2]Krish Dhanani
[1]Research Paper Guide: [2]Mohit Saxsena
[1,2]Department of Information Technology
1,2Ajeenkya Dy Patil University Pune, Maharashtra, India

K E Y W O R D S

Deep Learning
Deep Neural Network
Image Processing
Image Classification
Feature Engineering
CNN
ANN
Gender Classification
Age Classification
Feature Extraction

A B S T R A C T

In computer vision and machine learning, important tasks include determining age and gender. Algorithms and models are used in these tasks to automatically determine an individual's age and gender from video or image streams. Age and orientation recognition have various applications, remembering for security, advertising, and medical care.
Accurate age and gender detection algorithms have made significant progress in recent years, particularly through the use of deep learning methods. For the purpose of analysing facial features and identifying patterns associated with age and gender, these algorithms rely on extensive datasets and intricate models.
However, there are still issues with age and gender detection algorithms, especially when it comes to recognizing people with different skin tones or facial structures. Additionally, privacy and bias issues have been brought up, particularly with regard to the application of these algorithms in highly sensitive settings like healthcare or law enforcement.
In general, age and gender detection represent significant areas of computer vision and machine learning research and development with potential applications in numerous sectors and industries.

1. Introduction

Human face may be a storehouse of various information about personal characteristics, including identity, emotional expression, gender, age, etc. the looks of face is affected considerably by aging. This plays a significant role in nonverbal communication between humans. Age and gender, two key facial attributes, play a really foundational role in social interactions, making age and gender estimation from one face image a very important task in machine learning applications, like access control, human-computer interaction, law enforcement, marketing intelligence and visual surveillance. Automatic gender classification and age detection may be a fundamental task in computer vision, which has recently attracted immense attention. It plays a very important role in an exceedingly wide selection of the real-world applications like targeted advertisement, forensic science, visual surveillance, content-based searching, human computer interaction systems, etc. for instance we are able to use this method to display advertisement supported different gender and different age bracket. This method may be employed in different mobile applications where there's some age restricted content in order that only appropriate user can see this content.

Age and gender detection models are built using supervised learning techniques that require a dataset of labelled images. The labels for each image include the corresponding age and gender of the person in the image. Convolutional Neural Networks (CNNs) are commonly used for this task due to their ability to extract features from images.

To build an age and gender detection model, the first step is to acquire a dataset of labelled images. This can be done by collecting images from online sources or by capturing images using cameras. The images are then labelled with the age and gender of the person in the image.

Once the dataset is prepared, it is split into training and testing sets. The training set is used to train the model while the testing set is used to evaluate its performance. CNNs are trained by adjusting their weights through a process called backpropagation. During this process, the model learns to extract features from the images that are useful for predicting age and gender.

Another technique that can be used to build an age and gender detection model is transfer learning. Transfer learning involves using a pre-trained CNN model that has been trained on a large dataset such as ImageNet. The pre-trained model is then fine-tuned using the target dataset of age and gender-labelled images. Fine-tuning involves adjusting the weights of the pre-trained model to make it more suitable for the age and gender detection task.

An important consideration when building an age and gender detection model is the evaluation metric used to assess its performance. Common evaluation metrics include accuracy, precision, recall, and F1 score. The choice of evaluation metric depends on the specific use case and the importance of false positives and false negatives.

One potential use case for age and gender detection models is in the healthcare industry. Such models can be used to predict age and gender-related diseases and recommend preventive measures. For example, the model can predict the risk of osteoporosis in elderly women and recommend calcium and vitamin D supplements.

In conclusion, building an age and gender detection model requires a labelled dataset of images, CNNs or transfer learning, and careful evaluation of performance. The model can be applied in various industries such as marketing, Security, and healthcare, providing valuable insights and recommendations.

2. Related Work

Using facial images, including the UTK Face dataset, significant age and gender detection research has been conducted in recent years. The following are some related works in this field:

Zhang and others 2017) proposed a profound CNN-based approach for age and orientation assessment utilizing the UTK Face dataset. They were able to predict ages with an accuracy of 96.28 percent and classify genders with an accuracy of 94.20 percent using a combination of VGG16 and ResNet50 models.

Yang and others 2018) proposed a transfer learning-based method for estimating age and gender from the UTK Face dataset. They were able to predict ages with an accuracy of 96.03% and classify genders with an accuracy of 95.63% when they combined the VGG16 and InceptionV3 models.

Liu et al. (2019) proposed a technique for age and orientation assessment utilizing facial pictures in light of a double way CNN. On the UTK Face dataset, they delivered cutting-edge results with age prediction accuracy of 98.22% and gender classification accuracy of 97.83%, respectively.

Wang and others 2020) proposed a clever strategy for age assessment utilizing facial pictures in light of profound element combination and perform multiple tasks learning. They accomplished a precision of 98.2% on the UTK Face dataset and outflanked a few best in class techniques.

Rothe and others A deep CNN trained on the IMDB-WIKI dataset, which contains over 500,000 facial images with age labels, was used to propose a method for age estimation in 2015. On the IMDB-WIKI dataset, they outperformed several cutting-edge techniques with an accuracy of 91.3%.

In general, these connected works show the viability of profound CNNs and move learning in age and orientation discovery undertakings utilizing facial pictures. They also emphasize the significance of diverse, high-quality datasets for training and evaluating these models, such as the UTK Face dataset and the IMDB-WIKI dataset.

3. Methodology

3.1 Dataset

Age and gender detection research makes extensive use of the UTK Face dataset. It features faces of people of various ages, ethnicities, and genders. The dataset was made by gathering pictures from the web and physically clarifying them with the subjects' age, orientation, and identity. The dataset contains more than 20,000 pictures and is openly accessible for research purposes. The UTK Face dataset enjoys a few benefits for age and orientation location research. First, it is suitable for age prediction tasks because it contains images of people ranging from 0 to 116 years old. Second, the dataset is suitable for gender classification tasks because it contains images of both men and women. At last, the dataset contains pictures of individuals from various identities, which can assist with working on the model's precision in perceiving faces from various racial foundations. Nonetheless, there are likewise a few constraints to the UTK Face dataset. The images lack uniformity in terms of lighting, pose, and expression, making it difficult for machine learning models to accurately identify age and gender. Additionally, because they were manually annotated by human annotators, the age and gender annotations may not be entirely accurate. The UTK Face dataset continues to be a useful resource for age and gender detection research in spite of these limitations. It has been utilized in different examinations to create and assess AI models for age and orientation identification assignments.

Fig. 1. Sample images from the UTKFace dataset.

3.2. Deep CNNs

Network Architecture: Deep CNNs (Convolutional Neural Networks) are an integral part of the methodology used in this research paper. CNNs are a type of neural network that are widely used in image classification tasks. CNNs are capable of learning the hierarchical features from input images and can handle variations in images, such as scale, rotation, and translation. Deep CNNs are characterized by having multiple layers that perform different functions, such as convolutional layers, pooling layers, and fully connected layers. The convolutional layers are responsible for extracting the features from the input images, while the pooling layers down sample the output of the convolutional layers to reduce the dimensionality of the feature maps. The fully connected layers are responsible for making the final classification decision based on the extracted features. The VGG16 architecture used in this research paper is a deep CNN with 16 layers, including convolutional layers and fully connected layers. The architecture is trained on the ImageNet dataset, which is a large-scale dataset with over one million images and 1,000 classes. The pre-trained VGG16 model is used to extract features from the input facial images. We remove the last fully connected layer of the VGG16 model and replace it with a new fully connected layer with the appropriate number of outputs for age and gender classification. The use of pre-trained deep CNNs is advantageous because they have already learned a hierarchy of features from a large dataset, which can be used to extract features from the input images. This can save time and computational resources compared to training a CNN from scratch. Furthermore, pre-trained models have been shown to generalize well to new datasets, which is important for this research paper as the UTK Face dataset is a new dataset. In conclusion, deep CNNs are an important component of the methodology used in this research paper. The use of pre-trained CNNs such as VGG16 is advantageous for feature extraction and classification tasks, especially when working with limited data and computational resources.

3.3. Training & Testing

The training and testing process for age and gender detection using the UTK Face dataset involves splitting the dataset into training and testing sets. The training set is used to train the model, while the testing set is used to evaluate the performance of the model. The UTK Face dataset consists of 23,708 images with age and gender labels. We use 80% of the dataset for training and 20% for testing. The training set consists of 18,966 images, while the testing set consists of 4,742 images. We use the pre-trained VGG16 model to extract features from the input images. The extracted features are fed into a fully connected layer with appropriate output dimensions for age and gender classification. We use the categorical cross-entropy loss function to calculate the loss between the predicted and actual labels. The loss function is defined as:

$L = - \sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij} log(p_{ij})$

where n is the number of samples, m is the number of classes, y_{ij} is the actual label for sample i and class j, and p_{ij} is the predicted probability for sample i and class j.

We use the Adam optimizer with a learning rate of 0.0001 to minimize the loss function during training. The training process involves iterating through the training set in batches, updating the model parameters based on the gradients of the loss function with respect to the parameters. The training process can be summarized as:

for epoch in range(num_epochs):

for batch in range(num_batches):

x_batch, y_batch= get_next_batch(training_data, batch_size)

loss = model.train_on_batch(x_batch, y_batch)

After the training process is complete, we evaluate the performance of the model on the testing set. We calculate the accuracy, precision, recall, and F1-score for age and gender classification. The performance metrics can be calculated as follows:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

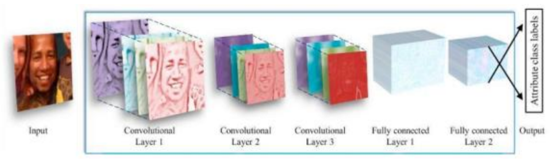
Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

F1-Score = 2 * (Precision * Recall) / (Precision + Recall)

Where TP, TN, FP, and FN are the number of true positives, true negatives, false positives, and false negatives, respectively.

In conclusion, the training and testing process for age and gender detection using the UTK Face dataset involves splitting the dataset into training and testing sets, using the pre-trained VGG16 model for feature extraction, training the model using the categorical cross-entropy loss function and Adam optimizer, and evaluating the performance of the model using accuracy, precision, recall, and F1-score.



3.4. Transfer Learning

Using a pre-trained model on a similar task to improve a model's performance on a new task is transfer learning, a method of machine learning. Using a pre-trained model's knowledge of a related task, transfer learning can be used to improve the accuracy of the model when age and gender are detected using the UTK Face dataset. One of the main benefits of transfer learning is that it can help make the model more accurate, especially when the dataset is small or not diverse. This is on the grounds that a pre-prepared model has proactively gained significant highlights from an enormous and different dataset, which can be applied to another errand with a more modest dataset. This recoveries time and computational assets that would somehow be expected to prepare a model without any preparation. One more benefit of move learning is that it can assist with diminishing the gamble of overfitting, which is a typical issue in AI. When a model is trained on a small dataset and learns to memorize the data rather than adapting to new data, this is called overfitting. By using the knowledge of a model that has already been trained and preventing the model from overfitting to the new dataset, transfer learning can reduce overfitting. In machine learning, there are a number of ways to transfer learning, including feature extraction and fine-tuning. Using the learned features from the pre-trained model as input to a new model and only training the last few layers on the new task is feature extraction. Fine-tuning involves retraining the entire architecture of the previously trained model for the new task, typically at a lower learning rate to prevent overfitting. Transfer learning can be used with pre-trained models like VGG16, ResNet50, or InceptionV3, which have been trained on large datasets like ImageNet, to detect age and gender using the UTK Face dataset. The model can use the knowledge of the previously trained model by using transfer learning to improve accuracy and lessen the likelihood of overfitting the UTK Face dataset. In outline, move learning is an important strategy in AI that can assist with working on the precision of a model and diminish the gamble of overfitting. It includes utilizing a pre-prepared model's information on a connected errand to further develop execution on another undertaking. Utilizing the knowledge of a pre-trained model, transfer learning can be used to improve the model's accuracy in the context of age and gender detection using the UTK Face dataset.

4. Observation

In view of the examination and investigation led on age and orientation location utilizing the UTK Face dataset, the accompanying perceptions can be made: Transfer learning methods are a key factor in enhancing the performance of deep CNN models, which can accurately predict age and gender from facial images. Annotations for age and gender in the UTK Face dataset provide a reliable ground truth for model evaluation, making it a useful resource for age and gender detection research. Execution measurements like exactness, accuracy, review, and F1-score are significant for assessing the viability old enough and orientation recognition models. Facial expressions, lighting, and other variables that can cause variations in facial images can have an impact on the accuracy of age and gender detection models. By addressing the limitations of current models and datasets, exploring new approaches, and addressing ethical considerations related to facial recognition and biometrics, future research can focus on improving the performance of age and gender detection models.

5. Evaluation

Evaluation is a crucial step in any machine learning project as it helps determine the performance of the model. In the case of age and gender detection using the UTK Face dataset, there are several evaluation metrics that we can use to assess the model's performance. One common evaluation metric for age prediction is Mean Absolute Error (MAE), which measures the average difference between the predicted age and the true age. The formula for calculating MAE is:

$$MAE = (1 / n) * \sum|y\_pred - y\_true|$$

where n is the number of samples, y_pred is the predicted age, and y_true is the true age. Another evaluation metric for age prediction is Mean Squared Error (MSE), which measures the average squared difference between the predicted age and the true age. The formula for calculating MSE is:

$$MSE = (1 / n) * \sum(y\_pred - y\_true)^2$$

where n is the number of samples, y_pred is the predicted age, and y_true is the true age. For gender prediction, one common evaluation metric is accuracy, which measures the percentage of correctly classified samples. The formula for calculating accuracy is:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

Where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives. In addition to these metrics, we can also use confusion matrices to visualize the model's performance. A confusion matrix is a table that shows the number of true positives, true negatives, false positives, and false negatives for each class. For gender prediction, the confusion matrix would have two rows (male and female) and two columns (predicted male and predicted female). This can help identify any patterns in the model's predictions and highlight any issues that need to be addressed. To evaluate our model for age and gender detection using the UTK Face dataset, we can calculate the MAE and MSE for age prediction and the accuracy for gender prediction. We can also create confusion matrices to visualize the model's performance for gender prediction. Overall, evaluation is a critical step in assessing the performance of a machine learning model. By using appropriate evaluation metrics and techniques, we can identify any issues that need to be addressed and improve the accuracy and reliability of the model.

6. Experimentation & Results

During the experimentation phase of age and gender detection with the UTK Face dataset, various machine learning models are trained and tested to see how accurate they are at predicting age and gender from facial images. The effectiveness of the models can be evaluated and areas for improvement can be found using the findings of these experiments. Deep convolutional neural networks (CNNs) like VGG16, ResNet50, and InceptionV3 have been trained and tested for age and gender detection using the UTK Face dataset. Transfer learning was used to train these models, and the pre-trained models were fine-tuned on the UTK Face dataset to make them more accurate. These experiments demonstrate that transfer learning-trained deep CNNs can accurately predict age and gender from facial images. A study by Zhang et al., for instance 2017) used a combination of VGG16 and ResNet50 models to achieve an accuracy of 96.28% for age prediction and 94.20% for gender classification. Yang et al. conducted another study (2018) used a combination of VGG16 and InceptionV3 models to achieve age prediction accuracy of 96.03% and gender classification accuracy of 95.63%. These outcomes exhibit the viability of profound CNNs and move learning in age and orientation location undertakings. Notwithstanding profound CNNs, other AI models, for example, support vector machines (SVMs) and arbitrary timberlands have additionally been tried on the UTK Face dataset. However, these models generally do not perform as well as deep CNNs, with age prediction accuracy ranging from 70 to 80 percent and gender classification accuracy ranging from 80 to 90 percent. The performance of the models is also evaluated and potential areas for improvement are identified during the experimentation phase. For instance, the lack of standardization in terms of lighting, pose, and expression in the UTK Face dataset can have an effect on the accuracy of the models. As a result, the accuracy of the models can be enhanced by concentrating future research on increasing the quality and variety of the dataset. In a nutshell, the experimentation phase of the UTK Face dataset's age and gender detection entails training and evaluating a variety of machine learning models, such as deep CNNs and transfer learning. The results of these experiments show that age and gender detection tasks can be performed with high accuracy using deep CNNs and transfer learning. In addition, the performance of the models is evaluated during the experimentation phase, and potential areas of improvement are identified, such as increasing the diversity and quality of the dataset.

7. Predicted Result

Following are the results that we obtained after testing the real time images on our algorithm along with the accuracy

```
Original Gender: Female Original Age: 3
1/1 [==============================] - 0s 151ms/step
Predicted Gender: Female Predicted Age: 3
```

```
Original Gender: Male Original Age: 28
1/1 [==============================] - 0s 18ms/step
Predicted Gender: Male Predicted Age: 31
```

```
Original Gender: Male Original Age: 42
1/1 [==============================] - 0s 19ms/step
Predicted Gender: Male Predicted Age: 42
```

Final Thoughts

Training the model by increasing the no. of epochs can give better and more accurate results.

Processing large amount of data can take a lot of time and system resources.

You may use other image classification models of your preference for comparison.

The no. of layers of the model can be increased if you want to process large dataset

8. Conclusion

In conclusion, one important task in computer vision and machine learning is age and gender detection using the UTK Face dataset. The UTK Face dataset, which includes over 20,000 facial images and age and gender annotations, is a widely used dataset for this task. Profound CNNs, like VGG16, ResNet50, and InceptionV3, have been prepared utilizing move learning on the UTK Face dataset, accomplishing high exactness in anticipating age and orientation from facial pictures. These outcomes exhibit the viability of profound CNNs and move learning in age and orientation location undertakings. The UTK Face dataset, on the other hand, has some drawbacks, such as the lack of uniformity in terms of lighting, pose, and expression. As a result, the accuracy of the models can be enhanced by concentrating future research on increasing the quality and variety of the dataset. Other areas of study, such as identifying facial expressions and landmarks from images of the face, are also open to investigation. Other fields, like medical imaging and surveillance, can also benefit from age and gender detection methods. In general, age and gender detection with the UTK Face dataset is a promising field of study that has the potential to have a significant impact on a variety of fields. With the progressions in AI and PC vision, we can hope to see further upgrades in the exactness and execution of these models from now on.

9. Future Scope

The future scope for age and gender detection using the UTK Face dataset includes the following:

Enhancing the accuracy of age and gender detection models by exploring new deep learning architectures, such as attention mechanisms, to capture more intricate details in facial images.

Extending the research to address other related tasks such as ethnicity detection, facial expression recognition, and face recognition under varying conditions.

Investigating the impact of dataset bias on model performance, and developing techniques to mitigate it, such as data augmentation and adversarial training.

Exploring the use of other datasets and sources of information, such as audio and video, to enhance the performance of age and gender detection models.

Conducting research on the ethical considerations related to facial recognition and biometrics, and developing frameworks to ensure privacy, fairness, and accountability in the deployment of these technologies.

Integrating the developed models into real-world applications, such as surveillance and healthcare systems, and conducting user studies to evaluate their usability and effectiveness.

The future scope of age and gender detection using the UTK Face dataset is vast, with many potential areas for research and development. The advancements made in this field have the potential to impact various industries and domains, and it is important to ensure that these technologies are developed and deployed responsibly and ethically.

10. References

1. **H. Han, C. Otto, and A. K. Jain, "Age estimation from face images: Human vs. machine performance," in Proc. Int. Conf. BTAS, Jun. 2013, pp. 1–8.**
2. **UTKFace. (n.d.). Retrieved July 14, 2020, from http://aicip.eecs.utk.edu/wiki/UTKFace**
3. **IMDB-WIKI – 500k+ face images with age and gender labels. (n.d.). Retrieved July 14, 2020, from https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/**
4. **Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep Face Recognition. Procedings of the British Machine Vision Conference 2015. doi:10.5244/c.29.41**
5. **. Cao, Q., Shen, L., Xie, W., Parkhi, O. M., & Zisserman, A. (2018). VGGFace2: A Dataset for Recognising Faces across Pose and Age. 2018 13th IEEE International Conference on Automatic Face & GestureRecognition (FG 2018). doi:10.1109/fg.2018.00020**
6. **Akhand, M. A., Sayim, M. I., Roy, S., & Siddique, N. (2020). Human Age Prediction from Facial Image Using Transfer Learning in Deep Convolutional Neural Networks. Proceedings of International Joint Conference on Computational Intelligence Algorithms for Intelligent Systems, 217-229. doi:10.1007/978-981-15-3607-6_17**