International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

FACIAL IMAGE-BASED AGE AND GENDER IDENTIFICATION USING DEEP LEARNING

[1] Dr. M. Sridhar, [2]Mr.P.S.S.K.Sarma, [3] Balina Shanmitha Jahnavi,[4]Gorripati Haritha,

[5]Gannina Harika,

[1] Professor, Rajamahendri Institute of Engineering & Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

[2] Assistant Professor,Dept.of CSE, Rajamahendri Institute of Engineering & Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

[3,4,5] Students, Dept.of CSE, Rajamahendri Institute of Engineering &Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

Abstract-
Every element of human contact relies on the ability to recognize people, from recognizing family members to recognizing coworkers. In this procedure, gender and age are essential factors. Computer scientists have seen an uptick in demand for face-recognition-based automated demographic analysis as AI becomes more pervasive in our daily lives. This research sets out to address that gap by doing population demographic analyses using face photos and making gender and age predictions. Topics covered in the research include estimating a person's age from a static face picture, gender categorization, and age classification. One approach makes use of transfer learning, while the other is based on deep Convolutional Neural Networks (CNNs). The second method involves taking a look at several backbone models to choose the best one for age and gender categorization. These models are VGG16, ResNet50V2, ResNet152V2, Xception, InceptionV3, MobileNetV3Small, and MobileNetV3Large. With meticulous research and analysis, this study adds to the progress of face recognition technology. It provides valuable insights for real-world applications in several fields including forensics, security surveillance, targeted marketing, and missing person identification.

Topics covered include facial pictures, age and gender categorization, age estimation using convolutional neural networks, and transfer learning.

I. INTRODUCTION

With the advent of smartphones and social media, people all over the world are capturing and sharing life's experiences at a rate never seen before, leading to a dramatic increase in the number of images created. The tendency is emphasized by data from Phototurial [1], which shows that 4.7 billion images are taken daily throughout the globe, demonstrating how ubiquitous photography is in modern life. On a daily basis, there are an astounding 5.0 billion images uploaded, which adds up to 1.81 trillion photos in total.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

This is an incredible rate of 57,000 photos each second. With an anticipated 2.3 trillion images taken yearly by 2030, this tendency is expected to only get worse. In the middle of this deluge of photos, selfies have exploded in popularity and now make up a significant percentage of all photographs shared online. Every every day, people capture an incredible 92 million selfies, which accounts for around 4% of all images captured.

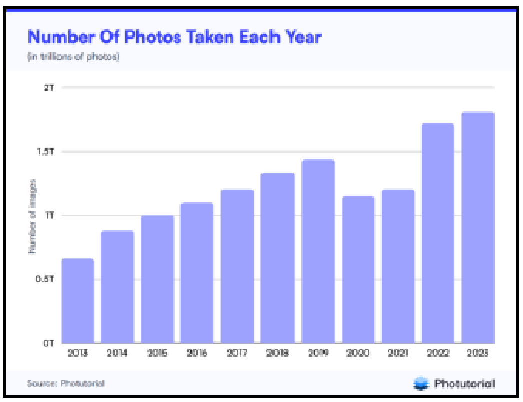

Fig.1. Number of photos captured annually (Photutorial/Matic Broz)

The increasing number of picture uploads in today's data-centric culture has led to a surge in interest in age and gender prediction within deep learning. While people are good at telling gender and identifying faces, it's far more difficult to put a precise age on someone. Consider that the mean absolute error (MAE) is the most used statistic to evaluate the accuracy of age predictions; this fact alone should serve to highlight the difficulty of the undertaking. According to studies, people may use an MAE between 7.2 and 7.4 (depending on the dataset parameters) to estimate the age of people beyond the age of 15 [2]. What this means is that the average human forecast is off by 7.2 to 7.4 years compared to the actual age. The development of deep learning and machine learning, however, opens the door of automating this procedure and, maybe, improving its accuracy. Facial aging is impacted by several variables, including genetics, lifestyle choices, expressions, and the environment [3]. Because of these different factors, people of the same age might seem quite different. Predicting age effectively is challenging due to its inherent complexity. Face aging is complicated and non-linear, which makes it difficult to carefully selecting datasets to use for gender and age prediction challenges. Regrettably, there is a dearth of datasets that sufficiently document the wide range of gender expressions and face aging trends. On top of that, most of the statistics out there are severely skewed in one direction or the other, with disproportionate numbers of people in certain age groups or belonging to certain genders [4][5]. Unreliable predictions and distorted model outputs might result from using such biased datasets. Hence, to improve the precision and consistency of gender and age prediction models, it is critical to combat dataset bias and work toward more balanced and representative datasets. Predicting an individual's age and gender from a photo of their face has several potential uses, each of which might provide new insights and useful solutions. Medical diagnostics and individualised treatment planning rely on accurate age estimate, while demographic analysis and patient classification benefit from gender prediction. Knowledge of consumer gender and age groups allows for more precise advertising and product suggestion strategies, which is a boon to the

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

retail and marketing industries. Furthermore, face recognition technology aids in the identification of suspects and the tracing of missing individuals in the security and law enforcement sectors. Algorithms that can anticipate a user's age and gender may improve their social media and entertainment platform experiences by providing more relevant ad suggestions and content. In the end, studies on determining a person's gender and age from face photos have the ability to greatly improve decision-making in many other fields. The following is a breakdown of the paper's parts. Section 2 offers a review of the relevant literature, and Section 3, which is further split into three sections, goes into depth into the technical issues. Section 4 details the outcomes, and Section 5 wraps up by outlining potential avenues for further study.

II.RELATED WORK

Scientists used anthropometric methods, which relied on ratios of different face feature measures, to forecast ages and genders in the beginning [3]. The dimensions of the eyes, nose, chin-to-forehead ratio, ear-to-ear ratio, and angles of inclination and distances between spots were all part of the set of measurements. Primary Component Analysis (PCA), Local Binary Patterns (LBP), Gabor filters, Linear Discriminant Analysis (LDA), and Scale-Invariant Feature Transform (SIFT) were among the first methods that relied on human feature extraction. Support Vector Machines (SVMs), decision trees, and logistic regression were among the conventional machine learning models that received these extracted characteristics. Researchers Geng et al. performed a research in 2007 that was one of the first to look at gender and age detection. To determine a person's gender and age, the authors of this study used a set of artificially constructed traits that included things like wrinkles, contour, and texture of the face. According to the results, these characteristics were reasonably accurate in determining gender and age. One major drawback of the research was that characteristics were hand-crafted, which might make them less applicable to other datasets.

In order to estimate ages, Hu et al. [6] used a method that included PCA, Support Vector Machines (SVM), and Uniform Local Binary Patterns (ULBP). However, a locally adjusted robust regression (LARR) approach was presented by Guo et al. [7] for age estimation. This technique combines SVM with Support Vector Regression (SVR). After using SVR to get a ballpark for the world's average ages, this technique switches to SVM for more accurate age estimates. One major problem with these methods is that it may be hard to get accurate anthropometric measurements, which makes it hard to generalize the model to other people of various ages and genders who could have comparable body types. There has been a significant uptick in the number of age and gender identification jobs using deep learning approaches. Using CNNs, or Convolutional Neural Networks, is a popular tactic. In their groundbreaking use of CNNs, Levi et al. [8] evaluated their performance using the Adience dataset, which includes gender and age categories. A

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

50.7% accuracy rate for age prediction and an 86.8% accuracy rate for gender categorization were produced by their inquiry. In particular, these results demonstrated how well CNNs perform when faced with facial images including age and gender information.

With the advent of D-CNN [9] for image classification tasks, the use of Convolutional Neural Networks (CNNs) for age and gender prediction got a major boost. With an emphasis on age classification using an ensemble of 20 networks on cropped faces from the IMDB-Wiki dataset, Rothe et al. [10] introduced DEX: Deep EXpectation of Apparent Age. One additional noteworthy work was that of Wang et al. [11], which fused characteristics generated from Deep CNN with Principal Component Analysis (PCA).

The development of computer vision neural networks has led to the emergence of several approaches. One of them is based on face recognition techniques and models, which suggests that there are strong generic models for faces that may be used for other tasks. Ordinal regression approaches used for age estimation are one area where wider models like VGG16 [13] are used as encoders in some investigations.

For the purpose of high-resolution gender and age classification in facial images, Zang et al. 2017 [14] presented ROR, a Residual network of Residual network. The model was pretrained on ImageNet, then fine-tuning on IMDB-WIKI 101, and further refinement on Adience. The model's accuracy for age prediction was 67.37%, which is lower than the outstanding gender categorization accuracy of 93.27%. Furthermore, processing rates were slower in the ROR model. Difficulty arose, for example, in accurately estimating age due to sensitivity to changes in alignment. In addition, general performance was impacted by restrictions in the Adience dataset, especially with respect to certain attributes. using a mean absolute error (MAE) of 4.1 years for age estimate and a 98.68% accuracy rate for gender prediction, VGGFace emerged as the clear winner in a 2018 study by Smith et al. [15] that investigated transfer learning using VGG19 and VGGFace models for gender and age recognition. The noise, mislabeling, and skewed age-gender distribution in the small dataset (MORPII) presented challenges. Notably, characteristics like hair length and head tilt were used for gender prediction. Three preprocessed picture streams and pretrained deep convolutional neural networks (DCNNs) were used in an ensemble learning strategy for face age estimate suggested by Yu et al. in 2019 [16]. The ensemble technique obtained an AEO of 88.20% and an exact match rate (AEM) of 45.57% using the IMBD-WIKI dataset. This demonstrated how important it is to have better algorithms that can handle a wider variety of data types and fast global search techniques. Difficulty in age estimate was caused by dataset difficulties, such as the time-consuming nature of DCNN for feature extraction and challenges with ambient light and complicated backdrops in IMBD-WIKI. In 2020, Olatunbosun et al. [17] tackled the computational and storage issues that come with bigger models by presenting a Lightweight Convolutional Neural Network (CNN) for actual and apparent age estimate.

The lightweight model increased its performance with MORP II, achieving Mean Absolute Errors (MAEs) of 3.05, 2.31, and 4.94 correspondingly by using datasets like FGNET and APPA-REAL. The significance of strong image processing algorithms, especially for raw photos, is highlighted by the model's effectiveness, especially in terms of decreased training time. The goal of introducing GRANET, a Gated Residual Attention Network, by Garain et al. in 2021 [18] was to address the shortcomings of earlier models by using several attention blocks. With the use of datasets like UTKFace, we were able to obtain an impressive age prediction accuracy of 93.7% and gender identification of 99.2%. Wikipedia recorded an MAE of 5.45 and AdienceDB an MAE of 10.57, although AFAD and FGNET both had MAEs of 3.10 and 3.23, respectively. The need for improved intelligence with partly obstructed pictures and the difficulty of classifying photos including children were obstacles. We were able to overcome obstacles caused by changes in lighting, position, and resolution by making use of the UTKFace dataset.

III.PROPOSED METHODOLOGY

A. Dataset

For this research, we used on the UTKFace dataset [19], which includes more than 20,000 cropped and aligned face photos with age, gender, and ethnicity labels. This collection includes 23,708 photos with varying degrees of lighting, position, resolution, occlusion, and face expression (not including six with missing age labels). The dataset's numerous picture attributes, including brightness, occlusion, and location, as well as its reasonably uniform distributions, led us to believe that it accurately represents the public. Figure 2 shows several examples of the UTKFace dataset. Each picture has a three-element tuple that says how old the person is (in years), what race they are (White-0, Black-1, Asian-2, Indian-3, and Others-4), and what gender they are (Male-0, Female-1).

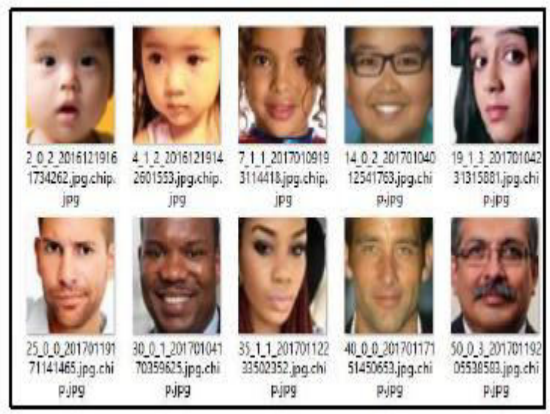

Fig.2. Sample images from UTKFace dataset

The gender, age, and ethnicity distributions of the dataset as examined by Exploratory Data Analysis (EDA) are shown in Figures 3, 4, and 5, respectively. Part B. of the deep A game-changer in visual data processing and understanding, Deep Convolutional Neural Networks (CNNs) have recently arisen as a critical technology in computer vision. These advanced neural networks use convolutional and pooling layers in a hierarchical structure to analyze and understand complicated pictures.

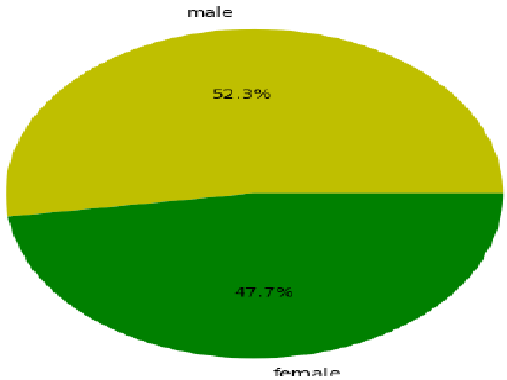

Fig.3. Gender distribution of UTKFace dataset

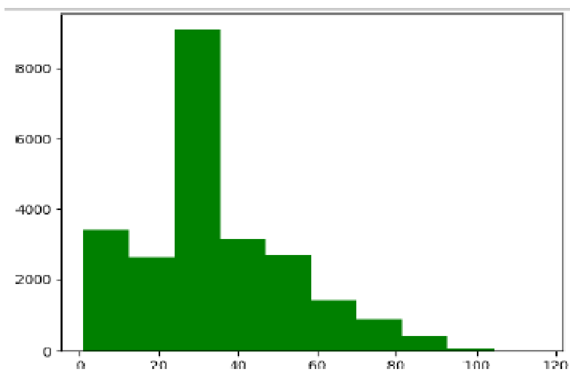

Fig.4. Age distribution of UTKFace dataset

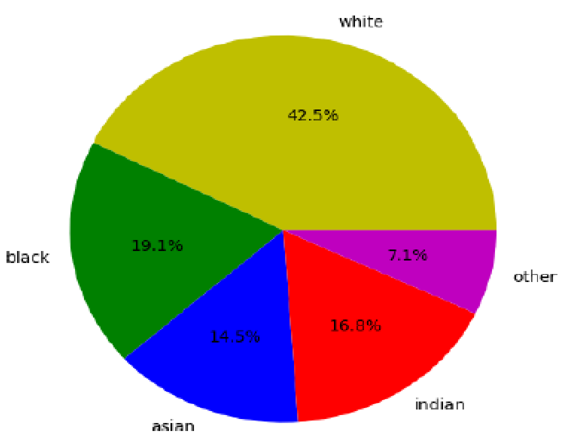

Fig.5. Ethnicity distribution of UTKFace dataset

Incredibly accurate object, face, and scene recognition is within the capabilities of deep convolutional neural networks (CNNs) thanks to their methodical feature and pattern extraction from raw image inputs. Their ability to capture both low-level aspects like textures and forms and structures, as well as high-level ideas like edges, makes them very successful for many image-related tasks. Deep convolutional neural networks (CNNs) have shown exceptional performance in a wide range of applications, including picture synthesis, object identification, semantic segmentation, and image classification. Applications like medical image analysis, autonomous driving, surveillance, and augmented reality greatly benefit from their capacity to automatically learn and adapt to diverse datasets. New developments in hardware and algorithms are driving research into deep convolutional neural networks (CNNs), which bodes well for future advances in AI and computer vision and promises to bring about an exciting new age of discovery and creativity.

IV.MODEL ARCHITECTURE

Overview

Age estimation, gender categorization, and deep Convolutional Neural Networks (CNNs) are the three main tasks that make up the underlying system. The approach divides people into five brackets based on their age: 0–24, 25–49, 50–74, 75–99, and 100–124. In order to correctly categorize people into these established age groups, CNNs use hierarchical layers of convolutional and pooling processes to extract subtle face information. Gender categorization works in a similar vein; it detects tiny facial traits that are exclusive to males and females. With the help of deep convolutional neural networks (CNNs), the system is able to produce impressive results when it comes to gender categorization. In addition, the algorithm uses face features to make accurate age predictions, treating age estimation as a regression job. In order to

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

capture subtle patterns and fluctuations linked to age, the system learns to map face characteristics to continuous age values using CNN architectures that are based on regression. Exhibiting outstanding performance across all three goals, this all-encompassing architecture deftly incorporates deep CNNs to tackle the varied difficulties of gender and age prediction tasks. The main tasks that the deep CNN technique tackles are summarized in Fig. 6.

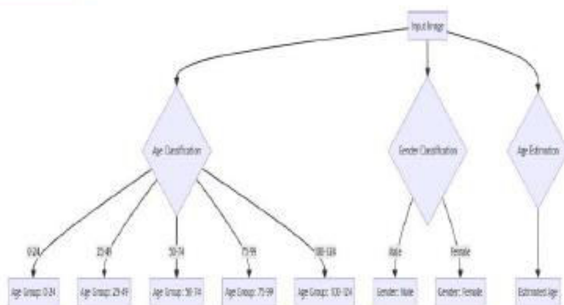

Fig.6. Core tasks addressed using deep CNNs

AGE AND GENDER CLASSIFICATION

Gender and age classification tasks used an architecture that was meticulously constructed to learn complex features and representations using a combination of convolutional and pooling layers. The model moved on to a succession of fully connected (FC) layers that were in charge of classification judgments after the first convolutional layers that extracted hierarchical features from the input RGB picture with dimensions (198, 198, 3). Gender categorization was reduced to a binary job, and age classification was organized into five separate groups to reflect different phases of human lifecycle. A multi-output classification method was used to concurrently handle the age and gender prediction goals. Using distinct FC layers for the two tasks, this method included a common feature extraction component for age and gender prediction. This shared feature extraction approach allowed the model to successfully capture underlying trends in age- and gender-relevant face features. The reasoning for this architectural decision was that there is a lot of overlap between the characteristics that are important for predicting age and gender, thus it would be more efficient to use a single feature extraction framework for both. To match the categorical character of the classification goals, categorical_crossentropy was created as the loss function for the age and gender classification tasks. To achieve a happy medium between computing efficiency and model stability, a batch size of 32 was used during training. With a starting rate of 0.008, the learning rate scheduler—implemented as a Keras callback function—dynamically altered the learning rate by half it every five epochs. Convergence was accomplished efficiently during training with the help of this adaptive learning rate method. There were two separate 22-epoch training cycles that made up the 44-epoch training program. Careful examination of the validation loss after the first training cycle ended showed signs of overfitting caused by the decreasing learning rate. After then, 22 further training epochs were run with the learning rate set at 0.002 to prevent overfitting and encourage further model improvement. A careful approach to training models and setting regularization parameters is required due to the persistence of overfitting, even with the introduction of regularization methods like Dropout and

Spatial Dropout.
In order to maximize the performance and stability of the model, layer-specific settings were painstakingly created. To improve local feature capture, 3x3 filters were used to convolutional layers. BatchNormalization was then applied to normalize activations of layers and increase training stability. After every convolutional layer, feature maps were downsampled using MaxPooling2D (2x2) operations. To avoid overfitting, a percentage of neurons were randomly deactivated during training using Dropout (0.2~0.3). To enhance regularization and reduce feature co-adaptation, Spatial Dropout (0.1~0.2) was added to convolutional layers with larger filter counts (128 or above). To make the model more capable of capturing complicated connections in the data, Rectified Linear Unit (ReLU) activation functions were used all across the network to induce non-linearity. A brief summary of the architectural configuration is given in Table I, which includes the kinds of layers, particular configurations, and the categories of output for jobs involving age and gender categorization.

Table I. Architectural Configuration of age and gender classification

Layer Type	Configuration
Input	RGB Image of size (198, 198, 3)
Convolutional	Filters: 3x3, Activation: ReLU
Batch Normalization	-
Max Pooling	Pool Size: 2x2
Dropout	Rate: 0.2~0.3
Convolutional	Filters: 3x3, Activation: ReLU
Batch Normalization	-
Max Pooling	Pool Size: 2x2
Dropout	Rate: 0.2~0.3
Convolutional	Filters: 3x3, Activation: ReLU
Batch Normalization	-
Max Pooling	Pool Size: 2x2
Dropout	Rate: 0.2~0.3
Fully Connected (FC)	Units: 128, Activation: ReLU
Dropout	Rate: 0.2~0.3
Fully Connected (FC)	Units: 64, Activation: ReLU
Dropout	Rate: 0.2~0.3
Age Classification	Output Categories: 0-24, 25-49, 50-74, 75-99, 100-124
Gender Classification	Output Categories: Male, Female

Fig.7 illustrates the architectural configuration of age and gender classification model.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362
IJARST

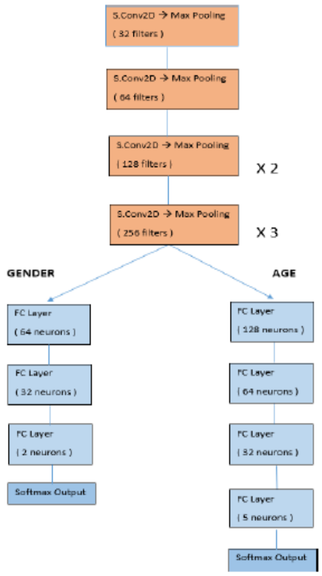

Fig.7. Model architecture of age and gender classification

AGE ESTIMATION

Careful consideration went into the development of the age estimation model in order to meet the complex issue of regression and achieve the goal of accurate age prediction from face photos. The model painstakingly scans RGB pictures using convolutional methods to extract prominent characteristics important for age prediction, operating on inputs scaled at 180 x 180 x 3. While training, the model makes use of the Adam optimizer to fine-tune its parameters according to the mean squared error (MSE) loss function, keeping an eye on performance measures like MAE and MSE. For flexible learning, a LearningRateSchedular dynamically adjusts the learning rate, which is set at 0.006 and then cut in half every 12 epochs for better model generalization and smoother convergence. There are a total of 95 epochs in the training process, split evenly into two halves of 45 epochs each. Importantly, after the first round of training, which yields slight performance improvements, another 50 epochs are painstakingly carried out to further hone the model's prediction abilities. Resetting the learning rate to 0.002 revs up the learning process and encourages faster convergence, allowing for this extension to run more smoothly. The model's architecture is based on convolutional layers that all use the same 3x3 filter size, with Dropout regularization (with values between 0.2 and 0.3) placed after each fully connected (FC) layer to further strengthen the model. Convolution layers with filters greater than 128 are reinforced with Spatial Dropout (varying from 0.1 to 0.2) to enhance regularization, and BatchNormalization is carefully added after each layer to stabilize and normalize activations. Rectified Linear Unit (ReLU) activation functions are used across all layers of the model, which helps in learning and extracting complex patterns related to age from face data. The architecture used for age estimate is shown in Table II, which provides a comprehensive overview. It shows the input size, loss function, optimizer, and training approach, among other things. Furthermore, the architectural diagram is shown in Fig. 8, which provides a thorough visual representation of the model's structure and components.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST ISSN: 2457-0362

Table II. Architectural Configuration of age estimation

Aspect	Details
Input Size	180 x 180 x 3
Loss Function	Mean Squared Error (MSE)
Optimizer	Adam
Metrics	Mean Squared Error (MSE), Mean Absolute Error (MAE)
Learning Rate Scheduler	Initial LR: 0.006. Halved LR every 12 epochs
Total Epochs	95 (45 + 50)
Training Strategy	Two-phase training: 1. Initial 45 epochs 2. Additional 50 epochs with LR adjustment
Filter Size	3x3
Regularization	Dropout (0.2~0.3) after every FC layer Batch Normalization after every layer. Spatial Dropout (0.1~0.2) after convolution layerswith filters > 128
Activation Function	Rectified Linear Unit (ReLU)

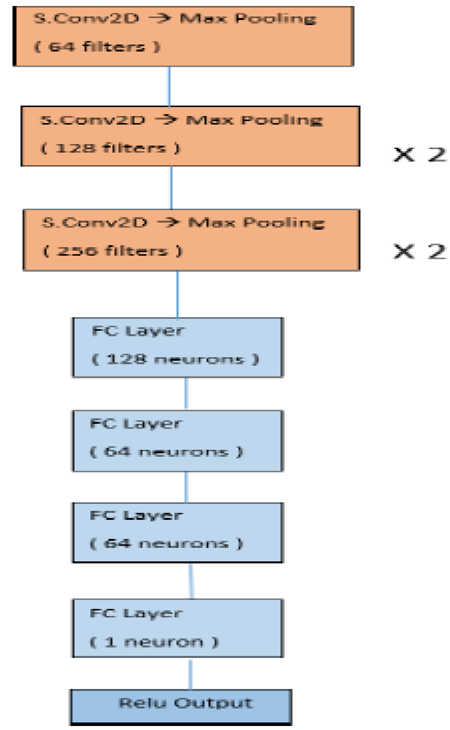

Fig.8. Model architecture of age estimation

C. Tranfer Learning

In machine learning and deep learning, a technique called transfer learning is used to improve the performance of another related but unrelated model by using information obtained from training the first model. Transfer learning is a subfield of deep learning that focuses on improving upon previously trained models using input from massive datasets like ImageNet. Training new models efficiently with little labeled data is made possible using transfer learning, which reuses learned feature representations from pre-trained models.

This reduces the requirement for substantial data annotation and processing resources. When there is a lack of labeled data or when it would be too computationally expensive to train a model from beginning, this method shines. Feature extraction, fine-tuning, and domain adaptation are just a few of the transfer learning methodologies that may be used to enhance the performance of a pre-trained model on a specific task. A fundamental method in deep learning research and practical applications spanning several fields including computer vision, natural language processing, and healthcare, transfer learning may speed up model training, improve generalization, and increase performance.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362
IJARST

OVERVIEW

Finding the best core model for age prediction and gender categorization is the goal of the transfer learning investigation. Xception, ResNet50V2, ResNet152V2, MobileNetV3Small, InceptionV3, and MobileNetV3Large are some of the possible models that will be evaluated in order to choose the best architecture to serve as the backbone of the system. By conducting a thorough comparison, several performance criteria including computing efficiency, accuracy, and resilience are used to assess the merits and shortcomings of each backbone model. The objective is to settle on a core design that can handle the difficulties of age prediction and gender categorization. Following this, the selected backbone will be fine-tuned using various methods, guaranteeing the creation of an efficient and optimum model suited to the current job needs.

FOUNDATIONAL ARCHITECTURES

The underlying architectures are described in depth in this section.
the VGG16 gene
The Visual Geometry Group (VGG) at the University of Oxford created convolutional neural network (CNN) designs, for which VGG16 is a cornerstone. Sixteen convolutional layers and a set of fully linked layers make up VGG16, a famously simple and effective network architecture. Many computer vision workloads like this architecture because to its simplicity and consistency. VGG16 is an excellent choice for feature extraction in transfer learning situations because of how well it captures fundamental picture characteristics using its stack of convolutional layers.
The RESNET50V6
ResNet50V2 is a member of the ResNet family, which included residual connections and hence transformed CNN architectures. ResNet50V2, a 50-layer network, uses residual connections to train deeper networks and avoid the vanishing gradient issue. Improved skip connections and normalization approaches are introduced by this design, which improves model performance and makes feature representation learning easier. ResNet50V2 is great at extracting complicated characteristics from picture data because of its depth and advanced architecture.
The RESNET152V2.
ResNet152V2 adds 152 layers to the ResNet architecture, making it more powerful and increasing the capacity of models. Training deep neural networks is made possible, similar to ResNet50V2, by using residual connections. Feature extraction and representation learning are two areas where ResNet152V2 shines because to its deep learning capabilities and parameter richness, which allow it to capture complex picture characteristics and patterns.
A rare instance
By including depthwise separable convolutions, Google's Xception proposes a radical break from conventional CNN designs. In order to make parameter learning and feature extraction more efficient, this design breaks down traditional convolutions into spatial and channel-wise processes. Xception is a great option for transfer learning applications, especially in situations

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

with limited resources, since it focuses on capturing fine-grained spatial correlations while keeping computation efficient.Section V The Inception module, which in InceptionV3 consists of parallel convolutional pathways with different kernel sizes, is another one of Google's creations. Strong representation learning is made possible by the network's ability to efficiently collect features at many sizes, according to its architecture. When it comes to jobs that need thorough feature extraction and semantic comprehension, InceptionV3 really shines when it comes to gathering various and multi-scale visual characteristics.

MINI MOBILENET V3 Designed for easy installation on embedded and mobile devices, MobileNetV3Small is a member of the MobileNet family. By using inverted residual blocks and depthwise separable convolutions, MobileNetV3Small finds a happy medium between model size and performance. It offers competitive performance in feature extraction and representation learning tasks despite its lightweight design and computational efficiency, making it suited for applications with limited computing resources.

Large MOBILENETV3

An expansion of MobileNetV3Small, MobileNetV3Large provides better speed but increases memory and computational complexity. The capabilities of MobileNetV3Large to capture patterns and represent features are improved with the inclusion of layers and parameters. Applications requiring more precision and advanced feature extraction capabilities may be met by MobileNetV3Large, which adheres to the efficiency principles of MobileNet architecture.

These models cover a wide variety of architectures, and each one has its own set of advantages and disadvantages. Our objective is to determine the best backbone architecture for our final model by comparing their performance on age prediction and gender categorization tasks. We want to optimize accuracy and computational efficiency.

TRAINING AND TESTING

Finding the best possible base architecture to support the final model's resilience is the main goal. The first step is to train a small network for a certain number of iterations in an iterative process. The study team is using this iterative training technique to find out which backbone architecture achieves robust performance criteria quicker and more significantly. The work aims to find the backbone architecture that is the most adaptable and performs the best across gender and age prediction tasks by methodically analyzing several designs.

Extensive testing has shown that optimising outcomes using a single model for age and gender predictions is not possible. Gender prediction accuracy is still below standard, in contrast to age prediction models, which demonstrate impressive levels of accuracy. This leads to a more sophisticated approach that calls for separate models to be used for gender and age forecasts. Each prediction job is unique in its complexity and difficulties, and this modeling bifurcation enables a specialized way to handle them. Researchers expect a considerable

International Journal For Advanced Research
In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362
IJARST

improvement in overall predicting performance by individually modifying structures and training procedures for each job. By deliberately dividing prediction tasks, we can better address the unique issues of age and gender prediction with an optimized strategy that takes into account their subtle complexities.

A Model for the Prediction of Age

Several backbone models, especially those associated with MobileNet, may be definitively ruled out as unfit competitors after analyzing the age model's learning curves in detail. This is because they regularly exhibit poor performance.

Because it consistently reduces training and validation loss, finally converges around a mean absolute error of 8.0, and stands out among the options, VGG16 is the most promising backbone. This steady pattern emphasizes its resilience, making VGG16 a good candidate for more evaluation.

Although Inception and Xception are alternative models with reduced training losses, they may be unstable because to the large difference between their validation and training curves. With its more consistent performance and flatter development curve, Inception stands out as a better alternative to Xception when it comes to balancing speed and resilience. The selection procedure is informed by this detailed examination, which guarantees conformity with the intended model attributes.

As a visual depiction of their performance measures, Fig. 9 shows the training loss and Fig. 10 shows the validation loss of all the age model's core architectures.

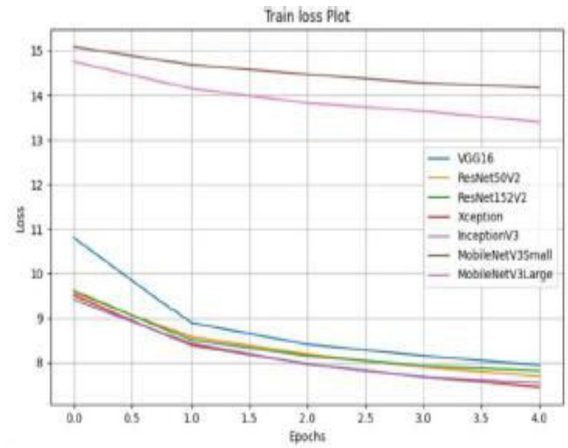

Fig.9. Training loss of all foundational architectures for age model

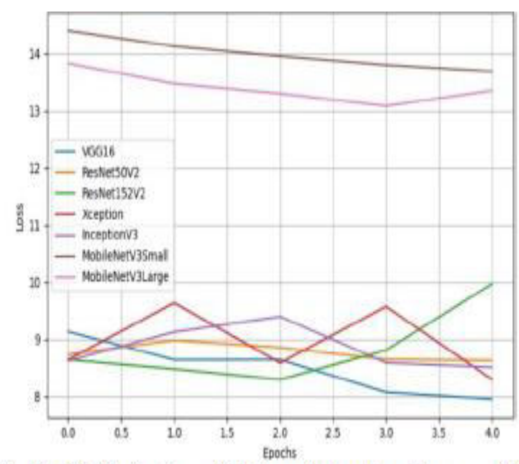

Fig.10. Validation loss of all foundational architectures for age model

GENDER CLASSIFICATION MODEL

The MobileNet models are eliminated from further consideration because to their persistent underperformance throughout the assessment of models for gender prediction accuracy.

By comparing the training and validation accuracy curves, the ResNet152V2 model stands out with its characteristic green curve, indicating its exceptional performance. For jobs requiring gender

IJARST
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

prediction, this model is an attractive option due to its top-notch training and validation accuracy.

The need for a more advanced baseline model in the final model development process is underscored by the large discrepancies between training and validation accuracy across all models. For this kind of model to adequately handle these differences, it has to have several layers.

With outstanding results in both training and validation, ResNet152V2 continues to be the best network architecture, even if a more refined baseline model is required. Although its validation accuracy is somewhat lower than that of ResNet152V2, ResNet50V2 nonetheless displays respectable training performance.

To help in the evaluation of their performance measures, Figures 11 and 12 showcase the training accuracy and validation accuracy, respectively, for all of the basic architectures that were investigated for the gender prediction job.

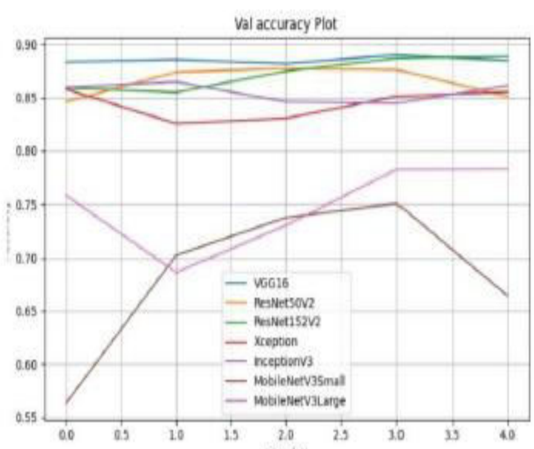

Fig.11. Training accuracy of all foundational architectures for gendermodel

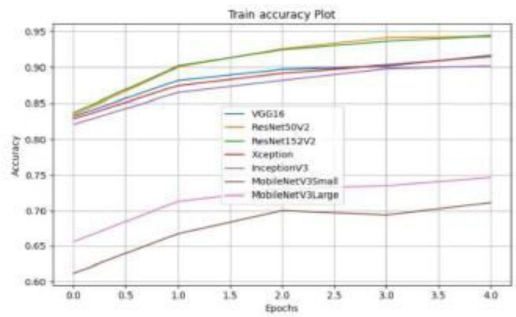

Fig.12. Validation accuracy of all foundational architectures for gender model

V.RESULTS AND DISCUSSIONS

An in-depth evaluation and explanation of the study's experimental results are provided in this part.

Deep CNNs Gender recognition, age classification, and age estimate are three basic tasks that are tackled in this research using deep convolutional neural networks (CNNs). The assessment scores for gender and age categorization are shown in increasing order in Table III, which also offers a thorough summary of the performance measures.

Table III. Results for age and gender classification

Train _loss	Train _age_ acc	Train _gen_ acc	Val _loss	Val_ age_ acc	Val_ gen_acc
1.836	0.662	0.909	2.124	0.632	0.874
0.748	0.911	0.958	1.617	0.704	0.879
0.934	0.836	0.932	1.619	0.719	0.880
0.052	0.991	0.991	1.116	0.741	0.897
0.786	0.839	0.993	0.903	0.741	0.885
0.138	0.965	0.994	0.749	0.827	0.950

Furthermore, the assessment scores for age estimate are shown in table IV, likewise in ascending performance order.

Table IV. Results for age estimation

Train_loss	Train_mae	Val_loss	Val_mae
177.7641	11.1566	199.3812	11.0324
17.6605	2.9786	83.6385	6.9328
89.6336	7.3960	66.4088	6.2225
58.5194	5.9128	30.8964	5.9402
44.6834	5.1108	44.0836	5.5679

IJARST
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

These tables are great for comparing how well different CNN architectures do on tasks like age estimation, gender identification, and categorization.

Predicting the Age of VGG16

After careful consideration, the study team settled on the VGG16 network as the Age Model's optimal backbone design. Consistently reducing training and validation losses and resulting in better model convergence were the deciding factors. In particular, the VGG16 network showed a validation loss of 7.56 and a training loss of 6.23. These measures show how well the network can learn and generalize from the training data while still performing well on validation data that it has never seen before. Figure 13 shows the findings graphically, with the Age Model's training and validation loss curves shown. The choice of VGG16 as the shared backbone highlights its effectiveness and dependability in enabling precise age estimates.

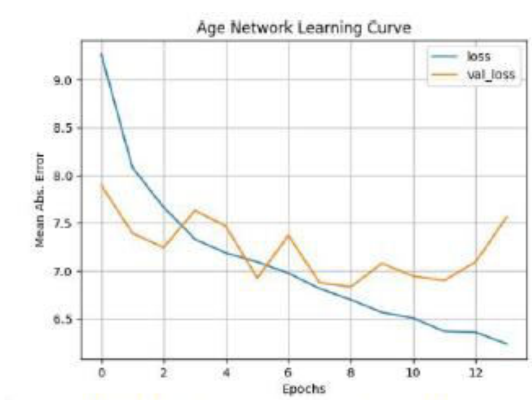

Fig.13. Loss curves of age model

ResNet152V2 Gender Classification

This model shows incredible robustness, even if it achieves somewhat lower accuracy. With a validation loss of around 0.30 and a training loss of close to 0.26, there is very little variation in the loss metrics. The accuracy metrics also show this little difference; the training accuracy is 88% and the validation accuracy is 87%. A 1% discrepancy between the two is the outcome of such consistency, demonstrating the model's remarkable dependability and stability.

Fig. 14 shows the gender prediction model's accuracy metrics, as well as the training and validation loss curves, which graphically reflect the results. With little variation in loss and accuracy values, this graphical depiction highlights the model's capacity to maintain balance between validation and training performance.

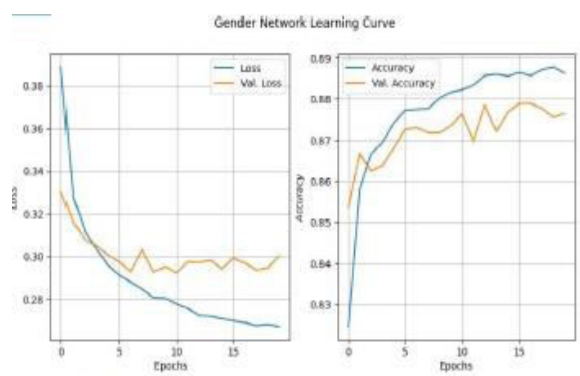

Fig.14. Accuracy and loss curves of gender model

VI.CONCLUSION

Finally, the results show that deep convolutional neural networks (CNNs) are quite good at detecting gender, classifying ages, and estimating ages. A variety of convolutional neural network (CNN) designs have been thoroughly tested and compared to see which ones work best for certain applications.

Using a larger and more representative dataset, we want to expand upon our previous findings in our future endeavors. Because of this, we will have a clearer picture of the biases that could influence our demographic analysis models. To lessen these biases, we want to test out several approaches. The more varied the data set, the better our models should be able to account for different types of users and scenarios. We will also investigate novel approaches, such as meta-learning and transfer learning, to enhance the robustness and flexibility of our models. We're looking forward to exploring various neural network types and experimenting with cutting-edge feature engineering techniques. Our mission is to significantly improve the accuracy and fairness of demographic analysis using face photos so that our models may be used in real-life scenarios by everyone. talent for managing complex data interactions. In the midst of a complicated global economic environment, our study equips decision-makers with useful insights, encouraging policy choices based on facts and better economic planning.

REFERENCES

[1] M. Broz, “How Many Photos Are There? 50+ Photos Statistics (2022),”photutorial.com,Feb.17,2022. https://photutorial.com/photos-statistics/

[2] H. Han, C. Otto, X. Liu, and A. K. Jain, “Demographic Estimation from Face Images: Human vs. Machine Performance,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 6, pp. 1148– 1161, Jun. 2015, doi: https://doi.org/10.1109/tpami.2014.2362759.

[3] R. Angulu, J. R. Tapamo, and A. O. Adewumi, “Age estimation via face images: a survey,” EURASIP Journal on Image and Video Processing, vol. 2018, no. 1, Jun. 2018, doi: https://doi.org/10.1186/s13640-018-0278-6.

[4] “IMDB-WIKI - 500k+ face images with age and gender labels,” Ethz.ch, 2015. https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/

[5] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, “VGGFace2: A Dataset for Recognising Faces across Pose and Age,” 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition(FG2018),May2018,doi: https://doi.org/10.1109/fg.2018.00020.

[6] L. Hu, Z. Li, and H. Liu, “Age Group Estimation on Single Face Image Using Blocking ULBP and SVM,” Proceedings of the 2015 Chinese Intelligent Automation Conference, pp. 431–438, 2015, doi: https://doi.org/10.1007/978-3-662-46469-4_46.

[7] G. Guo, Y. Fu, T. S. Huang, and C. R. Dyer, “Locally Adjusted Robust Regression for Human Age Estimation,” Jan. 2008, doi: https://doi.org/10.1109/wacv.2008.4544009.

[8] G. Levi and T. Hassner, “Age and Gender Classification using Convolutional Neural Networks.” Accessed: Jul. 13, 2023. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_workshops_2015/W08/p apers/Levi_Age_and_Gender_2015_CVPR_paper.pd

[9] Agrawal, K. K. ., P. . Sharma, G. . Kaur, S. . Keswani, R. . Rambabu, S. K. . Behra, K. . Tolani, and N. S. . Bhati. “Deep Learning-Enabled Image Segmentation for Precise Retinopathy Diagnosis”. *International Journal of Intelligent Systems*

and Applications in Engineering, vol. 12, no. 12s, Jan. 2024, pp. 567-74, https://ijisae.org/index.php/IJISAE/article/view/4541.[10] Samota, H. ., Sharma, S. ., Khan, H. ., Malathy, M. ., Singh, G. ., Surjeet, S. and Rambabu, R. . (2024) "A Novel Approach to Predicting Personality Behaviour from Social Media Data Using Deep Learning", *International Journal of Intelligent Systems and Applications in Engineering*, 12(15s), pp. 539–547. Available at: https://ijisae.org/index.php/IJISAE/article/view/4788

[11] O. Agbo-Ajala and S. Viriri, "Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces," The Scientific World Journal, vol. 2020, pp. 1–12, Apr. 2020, doi: https://doi.org/10.1155/2020/1289408.

[12] T. Kim, "Generalizing MLPs With Dropouts, Batch Normalization, and Skip Connections," arXiv.org, Aug. 22, 2021. https://arxiv.org/abs/2108.08186

[13] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv.org, Apr. 10, 2015. https://arxiv.org/abs/1409.1556

[14] K. Zhang et al., "Age Group and Gender Estimation in the Wild With Deep RoR Architecture," IEEE Access, vol. 5, pp. 22492–22503, Oct. 2017, doi: https://doi.org/10.1109/access.2017.2761849 .

[15] P. Smith and C. Chen, "Transfer Learning with Deep CNNs for Gender Recognition and Age Estimation," arXiv.org, Nov. 18, 2018. https://arxiv.org/abs/1811.07344 (accessed Mar. 17, 2024).

[16] N. Yu, L. Qian, Y. Huang, and Y. Wu, "Ensemble Learning for Facial Age Estimation Within Non-Ideal Facial Imagery," IEEE Access, vol. 7,pp. 97938–97948,2019 ,doi: https://doi.org/10.1109/access.2019.2928843 .

[17] O. Agbo-Ajala and S. Viriri, "A Lightweight Convolutional Neural Network for Real and Apparent Age Estimation in Unconstrained Face Images," IEEE Access, vol. 8, pp. 162800–162808, 2020, doi: https://doi.org/10.1109/access.2020.3022039 .

[18] A. Garain, B. Ray, P. K. Singh, A. Ahmadian, N. Senu, and R. Sarkar, "GRA_Net: A Deep Learning Model for Classification of Age and Gender From Facial Images," IEEE Access, vol. 9, pp. 85672–85689, 2021, doi: https://doi.org/10.1109/access.2021.3085971 .

[19] "Papers with Code - UTKFace Dataset," paperswithcode.com. https://paperswithcode.com/dataset/utkface