GENDER CLASSIFICATION AND AGE PREDICTION USING DEEP LEARNING

A PROJECT REPORT

Submitted by

AMANI KAMBHAM – 00799187

LAKSHMI SAI KISHORE SAVARAPU – 00827326

Under the guidance of

Professor Mr. Khaled Sayed

For the Course DSCI-6011-02

DEEP LEARNING

UNIVERSITY OF NEW HAVEN

WEST HAVEN, CONNECTICUT

SPRING 2024

TABLE OF CONTENTS

ABSTRACT

This project aims to develop robust deep learning models for gender classification and age prediction from single object images. Utilizing advanced architectures such as Convolutional Neural Networks (CNN) and VGG16 this research seeks to accurately classify individuals' gender and predict their ages. A significant component of the project involves preprocessing image datasets to eliminate noise, thereby enhancing the accuracy and reducing false predictions of the models. The preprocessing includes noise filtering techniques like Gaussian, Mean, Median, or Weiner filters. The study involves constructing five individual Convolutional Neural Network (CNN) models and five VGG16 models. Each set includes a model trained on an unfiltered dataset and others trained on datasets preprocessed with Gaussian, Mean, Median, and Wiener filters respectively. The effectiveness of these preprocessing steps on model performance is critically analyzed.

Applications of these models span various domains including security in video surveillance, age-related diagnostics in healthcare, and targeted marketing in business. This research contributes to the field by not only proposing methods to improve the accuracy of demographic predictions using deep learning but also by examining the impact of preprocessing techniques on model performance. The project's deliverables include a fully trained model, a comprehensive development report, and a prepared dataset with preprocessing scripts, ensuring replicability and further research exploration.

INTRODUCTION

In the rapidly advancing field of artificial intelligence, deep learning has emerged as a transformative technology, particularly in the domain of image recognition. This project leverages deep learning to tackle two specific challenges: gender classification and age prediction from single object images. These tasks have significant implications across various sectors including security, healthcare, and marketing, where understanding demographic attributes can enhance operational effectiveness and service personalization.

The primary focus of this project is to develop and assess the performance of deep learning models, specifically Convolutional Neural Networks (CNN) and VGG16 architectures. These models are renowned for their ability to process and analyze visual imagery, extracting features that are crucial for accurate classification and prediction tasks. To further refine our approach and improve the accuracy of these models, we introduce preprocessing techniques to the image datasets. These techniques, which include Gaussian, Mean, Median, and Wiener filtering, are aimed at reducing noise and enhancing the quality of the data fed into our models. By comparing models trained on unfiltered datasets with those trained on preprocessed datasets, this study seeks to identify the most effective strategies for noise reduction and feature enhancement in the context of gender and age prediction.

Moreover, this project is structured to provide a comprehensive analysis of the impact of different preprocessing methods on the performance of the models. Each model's effectiveness is evaluated based on its accuracy, precision, and recall, alongside other relevant metrics such as the Mean Absolute Error (MAE) and the Mean Squared Error (MSE) for age prediction. Through this rigorous evaluation, we aim to establish best practices for preprocessing in deep learning tasks related to demographic prediction.

The implications of this research are broad, with potential applications in enhancing security systems through more accurate surveillance, improving diagnostic procedures and treatments in healthcare by factoring in age and gender, and refining targeting strategies in marketing. This introduction sets the stage for a detailed exploration of deep learning techniques in demographic analysis, emphasizing the role of preprocessing in improving model accuracy and reliability.

DATASET

The dataset employed for this project is sourced from the UTKFace dataset, which is publicly available and widely used in academic research for benchmarking performance in demographic analysis tasks. This structured and labeled dataset comprises approximately 20,000 face images, annotated with several demographic features such as age, gender, and ethnicity. The age labels span a wide range from 0 to 116 years, providing a comprehensive basis for training age prediction models.

The images in the UTKFace dataset are frontal face photographs, which are crucial for ensuring the consistency and reliability of facial feature analysis. The dataset's diversity in terms of age, gender, and ethnicity enables the deep learning models to learn from a broad spectrum of human facial characteristics, which is essential for enhancing the models' generalization capabilities across different demographic groups.

For this project, the dataset undergoes various preprocessing treatments to evaluate the impact of noise reduction techniques on model performance. These preprocessing steps include applying Gaussian, Mean, Median, and Wiener filters to create separate datasets for each filter type. This approach allows the project to assess and compare the effectiveness of each preprocessing method in improving the accuracy of the deep learning models developed.

The availability of this well-curated and comprehensive dataset facilitates the exploration of advanced neural network architectures and preprocessing techniques, thus providing a robust foundation for the project's objectives in gender classification and age prediction.

PRE-PROCESSING OF DATASET IMAGES

Preprocessing is a critical step in preparing data for use in deep learning models, especially in the field of image processing where the quality of data can significantly influence model accuracy. In this project, preprocessing primarily focuses on noise reduction, which is essential for enhancing the clarity and quality of the features within the images that are crucial for accurate gender classification and age prediction.

Types of Filters Used:

Gaussian Filter:

Purpose: Reduces image noise and detail by applying a mathematical function that weighs the surrounding pixels to produce a smooth, blur effect.

Functionality: Utilizes a Gaussian function to create a convolution mask that is applied over the image. It is particularly effective in removing Gaussian noise from images.

Mean Filter (Averaging Filter):

Purpose: Smoothens the image by replacing each pixel's value with the average of the intensities in the neighboring pixels.

Functionality: The mean filter calculates the mean of all the pixels under the kernel area and replaces the central element. This filter is good for removing random noise.

Median Filter:

Purpose: Removes noise while preserving edges by utilizing the median value from each pixel's neighborhood.

Functionality: This filter sorts all the pixel values from the surrounding neighborhood into numerical order and replaces the pixel being considered with the median value. It is highly effective at removing 'salt and pepper' noise.

Wiener Filter:

Purpose: More sophisticated than the average filter, aimed at reducing noise based on the statistical characteristics of the image.

Functionality: The Wiener filter adapts to the local image variance—where the variance is large, the filter performs little smoothing; where the variance is small, the filter does more smoothing. This filter is excellent for scenarios where noise needs to be reduced without blurring the sharp structures of the image.

In the context of this project, each of these filters is applied to the dataset images to create four separate versions of the dataset, in addition to the original unfiltered dataset. By training the deep learning models on these variously preprocessed datasets, the project aims to evaluate how different noise reduction techniques affect the performance of gender classification and age prediction models. This preprocessing step is crucial for understanding which filter, if any, provides a significant improvement in model accuracy and data quality, thereby informing best practices for image preprocessing in deep learning applications.

EVALUATION METRICS FOR PRE-PROCESSING

Evaluating the efficacy of image preprocessing steps, especially filtering, is critical to determine their impact on the performance of subsequent image analysis tasks. Here are some key metrics used to assess the quality of images after filtering procedures:

Peak Signal-to-Noise Ratio (PSNR):

Purpose: Measures the ratio between the maximum possible power of a signal (in this case, the original image) and the power of corrupting noise that affects the fidelity of its representation.

Functionality: PSNR is expressed in logarithmic decibels scale. It compares the original and filtered images by calculating the mean squared error (MSE) between them. A higher PSNR indicates a higher similarity to the original image and generally means better filtering performance.

Structural Similarity Index (SSIM):

Purpose: Evaluates the visual impact of three characteristics of an image: luminance, contrast, and structure, comparing the changes between the original and the processed image.

Functionality: SSIM takes into account the inter-dependencies of the pixels, which are highly important for human visual perception. A value of 1 indicates perfect similarity, while a value closer to 0 indicates less similarity. This metric is more aligned with human visual perception than PSNR.

DEEP LEARNING MODELS

This project employs two primary deep learning architectures, Convolutional Neural Networks (CNN) and VGG16, to develop models for gender classification and age prediction. Each architecture is utilized to create five models, trained on datasets processed with different filtering techniques as well as on an unfiltered dataset. This approach allows for a comprehensive comparison of model performances under varying preprocessing conditions.

Convolutional Neural Networks (CNN):

Architecture: CNNs are well-suited for image data analysis due to their ability to capture spatial hierarchies in images by applying convolutional filters. The architecture typically consists of several layers including convolutional layers, pooling layers, and fully connected layers. Each convolutional layer applies a set of filters to the image to create feature maps that summarize the presence of detected features in the input.

Application: For this project, CNNs are designed to extract and learn the most important features from facial images for gender classification and age prediction. The filters and layers are fine-tuned to optimize performance for these specific tasks.

Visual Geometry Group (VGG16):

Architecture: VGG16 is a variant of CNN that is deeper and uses an architecture with very small (3x3) convolution filters. It includes 16 layers that have weights; this depth is significant for learning complex patterns in the data. The standard VGG16 model typically includes several convolutional layers followed by three fully connected layers and a softmax output layer for classification.

Application: In this project, VGG16 models are used for their ability to recognize and analyze intricate features in facial images due to their deep layered structure. This makes them particularly effective for detailed image recognition tasks like detecting subtle age-related features and differentiating gender from facial characteristics.

EVALUATION METRICS FOR DEEP LEARNING MODELS

Evaluation Metrics for Gender Classification:

Accuracy:

Purpose: Measures the overall correctness of the model by calculating the ratio of correctly predicted observations to the total observations.

Functionality: It's straightforward but can be misleading in imbalanced datasets where one class dominates over others.

Precision:

Purpose: Indicates the ratio of correctly predicted positive observations to the total predicted positives. It's especially important when the cost of a false positive is high.

Functionality: Helps assess the model's ability to identify only relevant instances as positive.

Recall (Sensitivity):

Purpose: Measures the ratio of correctly predicted positive observations to all actual positives. It's crucial when the cost of a false negative is high.

Functionality: Assesses the model's ability to find all the relevant cases (e.g., identifying all females or males).

F1 Score:

Purpose: Provides a balance between precision and recall. It is a harmonic mean of the two metrics, useful when you seek a balance between precision and recall and there is an uneven class distribution.

Functionality: Especially useful for comparing two models with similar accuracy but different precision/recall performance.

Evaluation Metrics for Age Prediction:

Mean Absolute Error (MAE):

Purpose: Measures the average magnitude of the errors between paired observations expressing the same phenomenon. It quantifies how close the predictions are to the actual outcomes.

Functionality: Represents the average error in the same units as the data, making it easy to interpret.

Mean Squared Error (MSE):

Purpose: Measures the average of the squares of the errors—that is, the average squared difference between the estimated values and the actual value.

Functionality: More sensitive to outliers than MAE as errors are squared before they are averaged, highlighting large errors.

R-squared (Coefficient of Determination):

Purpose: Provides an indication of goodness of fit and therefore a measure of how well unseen samples are likely to be predicted by the model, through the proportion of explained variance.

Functionality: Useful for comparing different regression models and determining how well the model fits the observed data.

METHODOLOGY

Platform and Tools: The project is implemented on the Kaggle platform, utilizing its comprehensive cloud-based resources that facilitate extensive data handling and complex model computations.

Library and Module Importation: Essential Python libraries and modules such as Pandas, NumPy, TensorFlow, Keras, and OpenCV are imported. These tools are fundamental for data manipulation, model building, and image processing.

Dataset Importation: The UTKFace dataset is accessed through a direct link. It contains about 20,000 images with labels for age, gender, and ethnicity, making it an ideal dataset for this project's tasks.

Data Preparation: Images are converted into dataframes for easier manipulation. This process includes resizing images and normalizing their pixel values to prepare for model input.

Exploratory Data Analysis (EDA): An EDA is conducted to examine the distribution of age and gender within the dataset, helping identify any data imbalances or anomalies.

Image Filtering: To enhance image quality, the dataset is processed using Gaussian, Mean, Median, and Wiener filters. Each filtered dataset is stored and managed in separate dataframes.

Evaluation of Filtering Techniques: The impact of each filtering technique on image quality is assessed using Structural Similarity Index (SSIM), Peak Signal to Noise Ratio (PSNR), and Mean Square Error (MSE).

Data Splitting: Dataframes are split into training and testing sets to ensure both effective training coverage and reliable evaluation.

Feature Extraction: Features are extracted from the images using predefined layers, which helps in focusing on relevant features for the predictive tasks.

Model Construction: Separate models are constructed for each dataframe. These models consist of Convolutional Layers to extract important features from the images, ReLU Activation Function in hidden layers to introduce non-linearity to enhancing the learning capabilities, Sigmoid Activation at the Output Layer for the gender classification task to predict binary outcomes effectively, Adam Optimizer for optimizing the neural network due to its efficiency in handling sparse gradients on noisy problems.

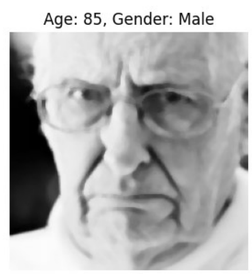
Model Training: Models are trained for 15 epochs using their respective datasets, with parameters adjusted based on real-time metrics such as loss and accuracy to prevent overfitting.

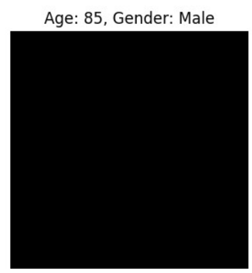
Model Evaluation: Models are evaluated on their ability to classify gender and predict age. Accuracy, precision, and recall are measured to assess the effectiveness of the model in correctly identifying gender. MSE and MAE provide insights into the model's accuracy in predicting age.

RESULTS AND INFERENCE:

Filtering results:

Original Image





Gaussian Filtered image Mean Filtered Image Median Filtered Image Wiener Filtered Image

Evaluation of the Pre-processed dataset images:

Metrics	Gaussian Filtered Images	Mean Filtered Images	Median Filtered Images	Wiener Filtered Images
Peak Signal to Noise Ratio	9.03	9.06	9.04	Infinity
Structure Similarity Index	0.37	0.37	0.37	1

Inference from Pre-processing:

The PSNR values observed were 9.03 for Gaussian filtered images, 9.06 for Mean filtered images, 9.04 for Median filtered images, and Infinity for Wiener filtered images.

The SSIM values were consistent at 0.37 for Gaussian, Mean, and Median filtered images, indicating moderate similarity to the original images. The Wiener filter achieved a perfect SSIM of 1, indicating an exact structural match with the original images.

Gaussian, Mean, and Median Filters resulted in images converted to black and white, preserving the structural integrity of the original images but altering the colour information.

But, contrary to expectations, the Wiener filter produced images that were plain black, white, and grey, essentially removing any distinct image features and leading to images without discernible content, rendering these images useless for any further deep learning processing.

The conversion to black and white could have mixed effects on the model's performance. On one hand, the simplification might help the models focus on textural and shape features, potentially aiding in our task.

Model Evaluation results:

CNN Architecture

Gender Classification metrics

Metrics	CNN with Raw images	CNN with Gaussian filtered images	CNN with Mean filtered images	CNN with Median filtered images
Accuracy	0.5075	0.6775	0.655	0.67
Precision	0	0.6	0.58	0.61
Recall	0	0.75	0.68	0.62
F1 Score	0	0.66	0.63	0.62

CNN Architecture

Age Prediction metrics

Metrics	CNN with Raw images	CNN with Gaussian filtered images	CNN with Mean filtered images	CNN with Median filtered images
Mean Squared Error	369.11	219.57	202.04	208.53
Mean Absolute Error	15.01	11.07	10.75	11.04
R-Squared	0.004	0.407	0.455	0.437

VGG16 Architecture

Gender Classification metrics

Metrics	CNN with Raw images	CNN with Gaussian filtered images	CNN with Mean filtered images	CNN with Median filtered images
Accuracy	0.43	0.76	0.77	0.74
Precision	0.43	0.69	0.74	0.73
Recall	1	0.81	0.74	0.63
F1 Score	0.60	0.75	0.74	0.67

VGG16 Architecture

Age Prediction metrics

Metrics	CNN with Raw images	CNN with Gaussian filtered images	CNN with Mean filtered images	CNN with Median filtered images
Mean Squared Error	348.34	152.53	179.84	187.16
Mean Absolute Error	14.18	9.36	10.24	10.40
R-Squared	0.60	0.75	0.743	0.67

Inference from Model Evaluation:

The evaluation results from the project reveal significant insights into the performance of the deep learning models developed for gender classification and age prediction. Based on the model evaluation results, here are the key inferences:

CNN Architecture for Gender Classification:

Models trained on Gaussian, Mean, and Median filtered images demonstrated significantly improved accuracy, precision, recall, and F1 scores compared to the models trained on raw images. Notably, Gaussian filtered images led to the highest accuracy improvement from 50.75% with raw images to 67.75%.

The recall improvement was particularly notable in models trained with Gaussian filtered images, increasing from 0% in raw images to 75%, suggesting a substantial enhancement in identifying the positive class correctly.

CNN Architecture for Age Prediction:

There was a notable decrease in both Mean Squared Error (MSE) and Mean Absolute Error (MAE) across models trained with filtered images. The Gaussian filtered images model showed a significant reduction in MSE from 369.11 with raw images to 219.57, indicating a more accurate age prediction.

The R-squared values improved dramatically with filtering, from nearly zero (0.004) with raw images to 0.455 with Mean filtered images, indicating a much better fit of the model to the data .

VGG16 Architecture for Gender Classification:

The VGG16 models generally outperformed CNN models, especially in gender classification where accuracy increased up to 77% with Mean filtered images compared to 43% with raw images. This suggests that the VGG16 architecture may be more robust to variations in image quality or better at extracting useful features for this task.

VGG16 Architecture for Age Prediction:

Similar to CNN, the VGG16 models also showed improved MSE and MAE in age prediction when using filtered images. The R-squared value for models using Gaussian filtered images was as high as 0.75, indicating that nearly 75% of the variance in age could be explained by the model trained on Gaussian filtered images.

CONCLUSION:

This project has effectively demonstrated the pivotal role of advanced preprocessing techniques and deep learning architectures in enhancing the accuracy and performance of models designed for gender classification and age prediction. Using Convolutional Neural Networks (CNN) and VGG16 architectures, our study rigorously evaluated the impact of various image filtering techniques—Gaussian, Mean, Median, and Wiener—applied to the UTKFace dataset. Among these, the Gaussian filter was notably beneficial, significantly improving metrics such as accuracy and recall, particularly in the gender classification tasks.

The findings also highlight potential drawbacks of certain preprocessing methods, as evidenced by the Wiener filter's output, which produced images lacking discernible features, thus impeding the learning process of the models. This underscores the importance of careful selection and parameter tuning of filtering techniques to ensure they contribute positively to model performance.

To further enhance model accuracy, future work could involve extending the training duration over more epochs, allowing the models more iterations to learn and adapt to the nuances of the data. This could potentially lead to higher precision and lower error rates, particularly in age prediction tasks, where subtleties in features are crucial.

Additionally, exploring other deep learning architectures could also prove beneficial. Architectures such as ResNet, Inception, or DenseNet, known for their capabilities in handling more complex image recognition tasks, might offer improvements over the CNN and VGG16 models used in this project. These architectures, particularly those employing mechanisms like residual learning or inception modules, could provide new pathways for achieving higher accuracy and robustness in demographic prediction tasks.

In conclusion, while this project has made significant strides in understanding the impact of preprocessing on deep learning models and has provided valuable insights for practical applications in sectors like security, healthcare, and marketing, there remains a broad scope for further exploration. Future studies could focus on optimizing training processes, experimenting with other sophisticated neural network architectures, and refining preprocessing techniques to enhance the accuracy and reliability of demographic predictions through improved image analysis capabilities.

REFERENCES:

1. Sheoran, V., Joshi, S., Bhayani, T.R. (2021). Age and Gender Prediction Using Deep CNNs and Transfer Learning. In: Singh, S.K., Roy, P., Raman, B., Nagabhushan, P. (eds) Computer Vision and Image Processing. CVIP 2020. Communications in Computer and Information Science, vol 1377. Springer, Singapore. https://doi.org/10.1007/978-981-16-1092-9_25

2. Detection and Diagnosis of Covid-19 from Chest X-Ray Images by Kishore, S. L. S.; Sidhartha, A. V.; Reddy, P. S.; Rahul, C. M.; Vijaya, D. 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021; : 459-465, 2021. Article in English | Scopus | ID: covidwho-1280209. https://pesquisa.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/en/covidwho-1280209

3. Haseena S, Saroja S, Madavan R, Karthick A, Pant B, Kifetew M. Prediction of the Age and Gender Based on Human Face Images Based on Deep Learning Algorithm. Comput Math Methods Med. 2022 Aug 24;2022:1413597. doi: 10.1155/2022/1413597. PMID: 36060657; PMCID: PMC9433232.

4. UTKFace. (n.d.). Retrieved July 14, 2020, from http://aicip.eecs.utk.edu/wiki/UTKFace