International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

Ageand Gender Classification Via Deep Learning Techniques

T.Vedasri[1], M Sanjay [2], K.Karthik[3],Dr.B.Laxmi Kantha [4]
[1,2,3] UG Scholar, Dept of IT, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
[4]Associate Professor, Dept of IT, St. Martin's Engineering College, Secunderabad, Telangana, India, 500100
vedaseisharma@gmail.com

Abstract:
Recognizing individuals is an innate and vital aspect of human interaction, from identifying loved ones to acquaintances in professional settings. Age and gender serve as fundamental attributes in this process. As artificial intelligence (AI) continues to integrateinto variousaspectsofourlives,the fieldofcomputer sciencehaswitnessed a surge indemand for automated demographic analysis based on facialrecognition. This paper aims to fulfill this need by analyzing the demographics of populations through facial images, predicting both age and gender. The study focuses on age classification, gender classification, and age estimation from static facial images. Two distinct methodologies are explored: one employing deep Convolutional Neural Networks (CNNs) and the other utilizing transfer learning. The latter approach entails an exploration of various backbone models, including VGG16, ResNet50V2, ResNet152V2,Xception,InceptionV3,MobileNetV3Small,andMobileNetV3Large,to determinethemostsuitablearchitectureforrobustageandgenderclassificationmodels. Through thorough investigation and analysis, this research contributes to the advancement of facial recognition technology, offering insights for practical applications in diverse domains such as forensics, missing person identification, personalized marketing, and security surveillance. Keywords—Age and gender classification, Age Estimation, Convolutional Neural Networks, Transfer Learning, Facial images.

1. INTRODUCTION

In the digital era, the widespread use of smartphones and social media platforms has sparked an unprecedented surge in image creation, as individuals globallycapture and share moments from their lives at an unparalleled pace. Data from Phototurial highlights this trend, revealing that an astounding 4.7 billion photos are taken worldwide everyday, showcasing the pervasiveness ofphotographyintoday's society. This staggering volume amounts to an annual total of 1.81 trillion photos globally, averaging an astonishing 57,000 photos per second or 5.0 billion per day. Projections suggest that by2030, this trend will continue to escalate, with an estimated 2.3trillion photos captured annually. Amidst this flood of images, selfies have become a ubiquitous phenomenon, constituting a notable portion of the global image-sharing landscape. An astounding 92 millionselfies are snapped daily, representing about 4% ofallphotostakeneachday. Age and gender predictionhas gained significant traction within the realm of deep learning, owing to the escalating influx of image uploads across the internet in today's data- centric society. While humans possess a natural proficiencyindiscerninggenderandrecognizingindividuals,accuratelyestimating age remains a formidable challenge. To underscore the complexity of this task, it is noteworthy to consider that

the most prevalent metric for assessing age prediction accuracy is the mean absolute error (MAE). Research indicates that humans can predict the age of individuals above 15 years with a MAE ranging from 7.2 to 7.4, depending onthe conditions of the dataset.This implies that, on average, humanpredictions exhibita deviation of7.2 to 7.4 years from the true age. However, the advent of machine learning and deep learning technologies presents an opportunity to automate this process and potentially achieve higher accuracy levels. It's important to recognize that facial aging is influenced not only by genetic factors but also by lifestyle choices, expressions, and environmental factors

2. LITERATURESURVEY

Literature Survey for Ageand Gender Classification Via Deep Learning Techniques

1. HowManyPhotosAreThere?50+PhotosStatistics

This article provides a comprehensive statistical analysis of the exponential growth of digitalimageryinrecent years.Itexamineshowthewidespreaduseofsmartphonesand socialmediaplatformshascontributed to a sharp rise in global photo production. Key statistics show the number of images uploaded daily and annually, presenting data from various sources to quantify the scale. It also discusses the increasing importance of image data in machine learning applications, such as facial recognition and age-gender classification, and the challenges ofprocessing such vast amounts ofdata. The article aims to inform readers about the growing importance of large-scale image databases in modern AI systems.

2. Demographic Estimation from Face Images: Human vs. Machine Performance

This paper investigates the performance differences between humans and machine learningalgorithms indemographicestimation fromfacial images. Theauthorscompare the accuracyofhuman observers with that ofstate-of-the-art algorithms for tasks such as age, gender, and ethnicity prediction. Through a detailed experimental setup, the paper highlights areas where machines outperform humans and vice versa, providing insight into the specific strengths and weaknesses of automated systems. It also introduces novel methods to improve machine performance in challenging scenarios, suchaslow-resolution imagesorocclusions. The studycontributestothe broader field of biometric authentication by improving demographic estimation systems.

3. AgeEstimationviaFaceImages

This comprehensive survey presents an overview of existing techniques for age estimationfromfacialimages.Thepapercategorizesmethodsintodifferentapproaches,suchas machine learning, deeplearning, andtraditionalcomputer visiontechniques. It provides a critical analysis of each approach's strengths and weaknesses, focusing on their performance inreal-worldapplications. Theauthorsdiscussvariousdatasetsused in age estimation research and highlight the importance of robust data preprocessing methods.Thepaperalsoexploresfuturetrendsinageestimation,suchasth eintegration of multimodal data and the potential for improving accuracy using advanced neural networks.

4. **IMDB-WIKI - 500k+ Face Images with Age and Gender Labels**Thisdataset,releasedbyETHZurich, isoneofthelargestpubliclyavailablecollections of face images labeled with age and gender information. The IMDB-WIKI dataset contains over 500,000 imagescollected from IMDB and Wikipedia. It is designed to facilitate researchonage and gender classification, offering a diverse range ofsubjects and varying image quality. The dataset has been extensively used in machine learning studies to train deep learning models for demographic estimation. Its large size and diversitymake it avaluable resource forimproving the robustness and generalizability of age and gender prediction algorithms.

5. VGGFace2: A Dataset for Recognising Faces across Pose and Age

The VGGFace2 dataset introduces a large-scale collection of face images designed to support research on facial recognition and demographic estimation. The dataset is notable for its inclusion of subjects across a wide range of poses, ages, and lighting conditions, making it ideal for training robust models that generalize well in the wild. Thispaperoutlinesthecreationofthedataset,thechallengesassociatedwith large-scale facedatacollection,anditsapplicationsinbothageandgenderclassification tasks.The authorsalsoprovidebenchmarksforpopular facialrecognitionalgorithmstestedonthe VGGFace2 dataset, emphasizing its utility in developing models capable of handling real-world variations.

3. PROPOSEDMETHODOLOGY

The proposed AI-driven approach to generating cybersecurity policies and procedures offers a range of significant advantages. Primarily, it enhances efficiency and speed by allowing the AI system to quickly process multiple datasets simultaneously, drastically reducing the time and effort required for policy creation. The use of standardized inputs ensures consistency and accuracy across generated policies, aligning them with industry best practices.

The use of standardized inputs ensures consistency and accuracy across generated policies, aligning them with industry best practices. Additionally, the system is adaptable and customizable, taking into account organization-specific parameters such as infrastructure, size, and culture, which enables it to produce tailored policies that meet the unique needs of each organization.This AI-driven system is scalable, easily accommodating policy generation for organizations of varying sizes and complexities, from small businesses to large enterprises with diverse technology stacks.

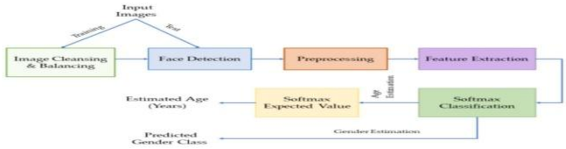
The proposed AI-driven approach to generating cybersecurity policies and procedures offers a range of significant advantages. Primarily, it enhances efficiency and speed by allowing the AI system to quickly process multiple datasets simultaneously, drastically reducing the time and effort required for policy creation. The use of standardized inputs ensures consistency and accuracy across generated policies, aligning them with industry best practices.

The use of standardized inputs ensures consistency and accuracy across generated policies, aligning them with industry best practices. Additionally, the system is adaptable and customizable, taking into account organization-specific parameters such as infrastructure, size, and culture, which enables it to produce tailored policies that meet the unique needs of each organization.This AI-driven system is scalable, easily accommodating policy generation for organizations of varying sizes and complexities, from small businesses to large enterprises with diverse technology stacks.

The proposed AI-driven approach to generating cybersecurity policies and procedures offers a range of significant advantages. Primarily, it enhances efficiency and speed by allowing the AI system to quickly process multiple datasets simultaneously, drastically reducing the time and effort required for policy creation. The use of standardized inputs ensures consistency and accuracy across generated policies, aligning them with industry best practices.

The use of standardized inputs ensures consistency and accuracy across generated policies, aligning them with industry best practices. Additionally, the system is adaptable and customizable, taking into account organization-specific parameters such as infrastructure, size, and culture, which enables it to produce tailored policies that meet the unique needs of each organization.This AI-driven system is scalable, easily accommodating policy generation for organizations of varying sizes and complexities, from small businesses to large enterprises with diverse technology stacks.

The proposed AI-driven approach to generating cybersecurity policies and procedures offers a range of significant advantages. Primarily, it enhances efficiency and speed by allowing the AI system to quickly



process multiple datasets simultaneously, drastically reducing the time and effort required for policy creation. The use of standardized inputs

Figure1.Proposed

ensures consistency and accuracy across generated policies, aligning them with industry best practices.The proposed AI-driven approach to generating of standardized inputs ensures consistency and accuracy across generated policies, aligning them with industry best practices. Additionally, the system is adaptable and customizable, taking into account organization-specific parameters such as infrastructure, size, and parameters such as infrastructure, size, and culture, which enables it to produce tailored policies that meet the unique needs of each organization. It also automates compliance by integrating the latest

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net
Vol-20 Issue-01 April 2025

industry standards (e.g., NIST, ISO) into the policy culture, which enables it to produce tailored policies that meet the unique needs of each organization.parameters such as infrastructure, size, and culture, which enables it to produce tailored policies that meet the unique needs of each organization.Cybersecurity policies and procedures offers a range of significant which enables it to produce tailored policies that meet the unique needs of each organization. It also automates compliance by integrating the latest industry standards (e.g., NIST, ISO) into the policy Primarily, it enhances efficiency and

The system follows a structured approach:

, AI system to quickly process multiple datasets simultaneously, drastically reducing the time and effort required for policy creation. The use of standardized inputs ensures consistency accuracy across generated policies, aligning them with industry best practices. Additionally, the system is adaptable and customizable, taking into account organization-specific parameters such as infrastructure, size, and culture, which enables it to produce tailored policies that meet the unique needs of each organization. It also automates compliance by integrating the latest industry standards (e.g., NIST, ISO) into the policy generation process, ensuring that the resulting documents adhere to current regulations. Finally, this AI-driven system is scalable, easily accommodating policy generation for organizations of varying sizes and complexities, from small businesses to large enterprises with diverse technology stacks.

Parallel processing of multiple datasets is a technique used to speed up data-intensive tasks by breaking down a large workload into smaller, independent sub-tasks that can be executed simultaneously on multiple processors or machines. This method is especially useful when processing large volumes of data or performing computationally intensive operations, such as data transformation, analysis, or machine learning model training. Let's elaborate on each step oftheParallel Processing of Multiple Datasetsprocess:

Step1.Preprocessing and Dataset Preparation

Step Description: Before parallel processing begins, it's essential to preprocess and prepare your datasets. This includes cleaning the data, handling missing values, transforming the data (e.g., normalizing, encoding), and splitting large datasets into manageable chunks for parallel execution.

Example: If you're processing sales data, preprocessing might involve:Removing duplicate entries.Filling in missing values or handling outliers.Splitting the data into subsets based on time periods, regions, or product categories.

Importance in Parallel Processing: Properly prepared data ensures that each parallel task gets a clean, consistent, and manageable subset of data, reducing the risk of errors during processing.

Step2.Define Task Granularity

Step Description: Task granularity refers to how finely tasks are divided for parallel execution. It determines whether the tasks will be small and fine-grained (many tasks) or large and coarse-grained (fewer tasks).

Example: If you're processing customer transactions:Fine-Grained Granularity: You might divide the data into smaller, transaction-level chunks and process each one independently across parallel tasks.Coarse-Grained Granularity: You could process transactions grouped by customer or by region, resulting in fewer, larger chunks.

Importance in Parallel Processing: Choosing the appropriate task granularity is critical for balancing the workload across

speed by allowing the It also automates compliance by integrating the latest industry standards (e.g., NIST, ISO) into the policy This AI-driven system is scalable, easily accommodating policy generation for organizations of varying sizes and complexities, from small businesses to large enterprises with diverse technology stacksthat meet advantages.parameters such as infrastructure, size, and culture the unique needs of each organization. It also automates compliance by integrating the latest industry standards (e.g., NIST, ISO) into the policy.

available processors. Fine-grained tasks might introduce high overhead due to excessive task management, while coarse-grained tasks might not fully utilize available resources. Proper granularity ensures efficient parallel processing without wasting resources.

Step3.Parallelization Strategy

Step Description: Once the granularity of the tasks is defined, the parallelization strategy determines how the tasks will be executed in parallel. It involves selecting the parallelization method (e.g., data parallelism, task parallelism) and how tasks will be distributed across available processors or nodes.

Example: If you're analyzing data from a large retail database:

Data Parallelism: You could distribute subsets of the data (e.g., by store or time period) across multiple processing units, where

Task Parallelism: Different tasks, such as data preprocessing, feature extraction, and model training, could be executed concurrently on separate processors or machines.

Importance in Parallel Processing: The parallelization strategy determines how efficiently tasks are executed concurrently. A well-defined strategy maximizes resource utilization and minimizes delays due to task dependencies. A poorly designed strategy can lead to imbalance, where some processors are idle while others are overloaded.

Step4.Parallel Execution

Step Description: Parallel execution refers to the actual running of tasks concurrently. This step involves distributing the defined tasks across available computing units (e.g., CPU cores, nodes, GPUs) and ensuring that each unit works on its assigned task without unnecessary delays.

Example: For a machine learning application:Distributed Execution: If using a distributed system like Apache Spark, tasks can be distributed across different nodes in a cluster to process large datasets in parallel.Multi-core Execution: On a single machine, tasks might be distributed across multiple CPU cores or GPUs for faster computation.

Importance in Parallel Processing: Efficient parallel execution ensures that tasks are processed concurrently, reducing the overall computation time. It's crucial for maximizing the utilization of available hardware, such as multi-core processors or distributed computing environments, thereby performance.

Ste5.Handling Data Synchronization and Communication

Step Description: When tasks are executed in parallel, they often need to communicate and synchronize their progress, especially in distributed systems. This step involves ensuring that data consistency is maintained, tasks that depend on each other can exchange information, and any shared resources are handled correctly.

Example: If you're processing customer feedback data:

Synchronization: Ensuring that one task does not overwrite or conflict with another task when updating shared resources (e.g., aggregated statistics).

Communication: In a distributed setup, nodes might need to share intermediate results, like sending partial sums back to a central node for final aggregation.

Importance in Parallel Processing: Synchronization and communication are crucial to avoid data inconsistencies and race conditions. Without proper synchronization, tasks might overwrite results, or some tasks might complete before others, leading to

International journal of imaging science and engineering

ISSN: 1934--9955 www.ijise.net

Vol-20 Issue-01 April 2025

incorrect outputs. Efficient communication ensures that tasks

work together seamlessly.

Step6.Post-Processing and Result Integration

Step Description: After parallel tasks are completed, the results need to be aggregated, combined, and possibly further processed to generate the final output. This step ensures that the individual outputs of parallel tasks are merged correctly and useful insights are extracted.

Example: If you processed sales data by region in parallel:

Merging Results: Combine the processed sales statistics (e.g., total sales, average sales per region) from each parallel task.

Post-Processing: You might apply final transformations, like aggregating regional sales to generate a nationwide total or calculating performance metrics like sales growth.

Importance in Parallel Processing: Post-processing and result integration ensure that the output of parallel tasks is meaningful and accurate. Without careful result merging, partial results might not align correctly, leading to errors in the final output.

Step7.Error Handling and Fault Tolerance

Step Description: Error handling and fault tolerance ensure that the parallel processing system can recover from failures, such as task crashes or resource unavailability. This step involves detecting errors, managing task retries, and handling failures gracefully.

Example: In a large distributed system:Error Detection: If a node fails to process its assigned data, the system detects the failure and marks the task as incomplete.Fault Tolerance: The system can reschedule the failed task on another node or retry it from a checkpoint.**Importance in Parallel Processing:** Robust error handling and fault tolerance prevent data loss or corruption in the event of hardware or software failures. This ensures that parallel processing can continue without significant

.Step8.Output and Cleanup

Step Description: After parallel tasks are completed, the results are written to thedesired output location, and any resources used during execution are cleaned up. This includes closing files, freeing memory, and releasing other system resources.

Example: For a data analysis task:

Output: The final results of the parallel tasks (e.g., cleaned datasets, statistical analysis) are written to output files or databases.

Cleanup: Unused resources such as memory, threads, or temporary files are properly released to avoid system resource leaks.

Importance in Parallel Processing: Output and cleanup ensure that resources are efficiently managed and that no lingering processes or files cause issues. Proper cleanup also avoids memory leaks and ensures that the system remains responsive for future tasks. Additionally, ensuring output integrity guarantees that results are accessible and correct.

4.EXPERIMENTAL ANALYSIS

Figure 2: Home Page

Figure3: User Login

Figure 4: Admin Page

The projectsays Thesemodelsrequiredexplicithuman interventiontoextractrelevantfeaturesfromimages,suchasfaciallandmar ksor textures,beforemakingpredictions.Whilethesemethodsworkedwellfors maller datasets,theyoftenstruggledwiththecomplexityoflarger,morediverseim age datasetsandhad limitations in accuracydue to their reliance on hand-crafted features. Withtheadvent ofdeep learning, specificallyConvolutionalNeuralNetworks(CNNs), the performance of age and gender classification systems improved drastically.

5. CONCLUSION

In conclusion, this study has demonstrated the effectiveness of deep convolutional neural networks (CNNs) in addressing gender detection, age classification, and age estimationtasks.Throughcomprehensive evaluationand comparison, the performance of different CNN architectures has been assessed, providing valuable insights intotheir suitability for these tasks. In our future work, we plan to build on what we've learned so far by using a bigger and more balanced set of data. This will help us understand betterthe biases that can affect our models when analyzing demographics. We want to try out different methods to reduce these biases. By having more diverse data, wehope our models will work better for all kinds ofpeople and situations. We'll also explore new techniques like transfer learning and meta-learning to make our models stronger and more adaptable. We're excited to try out advanced ways of engineering features and testing different types of neural networks. Our goal is tomake big strides in how we analyze demographics from facial images, making our models more accurate and fair for everyone to use in real-life situations. aptitude for handling intricate data relationships.

REFERENCES

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

[2] H.Han,C.Otto,X.Liu,andA.K.Jain, "DemographicEstimationfromFaceImages: 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. [Online]. Available: https://data.vision.ee.ethz.ch/cvl/rrothe/imdbwiki/

[5] Q.Cao,L.Shen,W.Xie,O.M.Parkhi,andA.Zisserman,"VGGFace2: ADataset for Recognising Faces across Pose and Age," *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, May 2018, 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 BlockingULBPandSVM,"*Proceedingsofthe2015ChineseIntelligentAutomation 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,andC.R.Dyer,"LocallyAdjustedRobust Regressionfor HumanAgeEstimation,"Jan.2008,doi: https://doi.org/10.1109/wacv.2008.4544009.

[8] G. Leviand T. Hassner, "Age and Gender Classificationusing ConvolutionalNeural Networks," *CVPR Workshops 2015*. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_workshops_2015/W08/papers/Levi_Age_ and_Gender_2015_CVPR_paper.pdf.

[9] A. Krizhevsky, I.Sutskever, and G. E. Hinton, "ImageNet Classification with Deep ConvolutionalNeuralNetworks," *Communicationsof theACM*,vol. 60, no. 6,pp. 84– 90, May 2012, doi: https://doi.org/10.1145/3065386.

[10] R. Rothe, R. Timofte, and L. Van Gool, "Deep Expectation of Realand Apparent Age froma Single Image Without Facial Landmarks," *International Journal of Computer Vision*, vol. 126, no. 2–4, pp. 144–157, Aug. 2016, doi: https://doi.org/10.1007/s11263-016-0940-3.