JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

CALCULATION FACIAL AGING PATTERN USING DIGITAL FORENSIC SCIENCE AND MACHINE LEARNING

[1] Dr.Yazdani Hasan, [2] Sumbul Afroz, [3]Deepak Kumar
[1] Associate Professor, [2] Assistant Professor,[3]Associate Professor
[1] , School of Engineering and Technology(CS), [2,3] School of Engineering and Technology(EC)
[1,2,3] IIMT University, Meerut, India

ABSTRACT

To identify perpetrators and victims of both online and offline crimes, police agencies frequently employ closed circuit television, mobile photos and videos, open-source intelligence (i.e., social media/web data mining), and other photographic evidence sources. Police officers and witnesses frequently include human characteristics, such as age, height, weight, gender, hair color, etc., in their descriptions of unidentified suspects. In some cases, such as in child abuse investigations, the age of the victim might influence how the crime is classified. In order to detect soft biometric features like age and gender in digital photos and help detectives and investigators move their cases forward, a number of automated machine learning-based approaches have been devised.

The assessment of current cognitive age prediction services is presented in this research. To find trends and problems with the various services' performance, an evaluation and comparison of those services was done. The noteworthy absence of sufficient sample photos in particular age groups, i.e., young children and the elderly, is one of the major contributing factors hindering the accurate development of the services under investigation. A dataset generator was created as a solution to this problem. It uses collections of several imbalanced datasets to create a balanced, curated dataset of digital photos that are tagged with the correct age and gender.

KEYWORDS

Artificial Neural Network , Digital forensic , Dataset, Transfer Learning

1. INTRODUCTION

While it just takes a fast glance at a digital image for a person to detect and identify many characteristics like age, gender, ethnicity, and facial expressions, robots must undergo extensive training in order to recognize characteristics in images. With the launching of Apple's iPhone X, facial recognition has lately made a comeback in the popular media. It has been the key feature of various devices over the past couple of years. This phone replaces the conventional fingerprint authentication method with face detection technology as its primary unlocking and authentication security feature.

Hi-tech Around the world, there is ongoing development of face recognition for many different uses. Face recognition technology has been employed in China for a variety of purposes, including ride hailing service driver identification and payment.with a smile, identification of the jaywalker, etc. . Numerous face recognition software programmed have already entered the market; their capabilities and scope sometimes astound clients. A variety of offences benefit from correct age assessment, which has applications beyond child abuse investigations.
In our environment, age estimate products are getting increasingly prevalent. The emergence of pervasive

smart face detectors will have a significant influence on users' and customers' lives. By switching to a face recognition-based authentication element, system security may be strengthened to thwart impersonation attempts. Identity theft, fake exam takers, and card cloning may be prevented through facial security checks. Age estimate is also useful for adult entertainment venue entry, the purchase of age-restricted commodities like alcohol and cigarettes, age-targeted advertising, and, more lately, the identification of Syrian refugee children via services like how-old.net. The situations stated above are only a few instances of the numerous applications that "multi-layered deep learning technology for highly intelligent systems" might be used for. It has always been difficult to determine a subject's age accurately for study in a variety of professions. From counting the yearly rings of wood growth to establish a tree's age to gauging a bone's skeletal maturity to establish a living thing's age. A number of techniques, including measurement-driven anthropomorphic analysis and the use of machine learning algorithms, have been used to estimate age, and the accuracy of the results is continually improving. Furthermore, it is a research topic since it has been shown that the methods now in use are insufficiently trustworthy to execute the task of separating the marginally underage from the marginally overage. Describes a compelling use for face recognition in photo organization that is used in digital forensics and cybercrime investigation. A time-consuming, highly-skilled digital forensic examination must be performed in order to process the sizeable quantity of data that is often seized and processed. Additionally, because of the necessity for speedy processing, confiscated devices are frequently of information to further a probe into what may be a life-or-death situation.

Forensic investigators have long found it laborious to search a disk's surface for information with potential probability. However, cutting-edge machine learning and computer vision algorithms (central components of artificial intelligence) can assist digital forensic professionals in carrying out automatic file classification, flagging various sorts of material as significant, and more. to a case in court, and lessen the exposure of child abuse material in digital forensic laboratories.

Investigations into child abuse are frequent occurrences in law enforcement organizations all around the world. Due to the increasing use of anonymization technologies, private P2P networks, and cloud-based KVM systems, these investigations have become a difficult undertaking. The fight against human trafficking and the reduction of child abuse material (CAM1) have been top priorities for law enforcement and child protection communities across the world. The deep web, which has developed into a center for illegal black markets and is where organized cybercrime organizations operate, is where pedophiles may exchange large amounts of CAM. They frequently do this to fit in with a community of other pedophiles and eventually acquire access to additional repositories of illegal content [7].

1.1. CONTRIBUTION OF THIS WORK

This work's contribution may be summed up as follows: • An review of current facial recognition techniques.

. CAM is the term used by and, despite various nomenclatures being interchangeably used in the literature for this unlawful substance.

• Performance assessment of offline and cloud-based face recognition models for age estimate accuracy and the impact of the subject's gender and actual age on the algorithms' predictions.

• The introduction of a tool made to produce distinct, evenly distributed random pictures by age and gender from several face image datasets, including FG-NET, FERET, IMDB-WIKI, MEDS, and YFCC100M.

In order to save time and money, it is crucial for the research community to share datasets. To train machine learning models and forecast more information from incoming data, facial picture datasets tagged with age and gender are required. emphasizes the value of dataset sharing among the scientific community in favour of reproducing findings.

Datasets from earlier work on age estimate have been made available. [14] acknowledged the lack of facial data and provided a dataset that gathered real-world photographs from open-source Flickr albums to address this shortcoming. Eight age groups were specified after the age and gender were noted by observation.

In order to create a collection of photos tagged with true age and gender, [9] crawled more than 500,000 celebrity photographs from Wikimedia and IMDB in 2016. With a convolutional neural network (CNN) with 16 weight layers that was pre-trained on ImageNet (an image dataset that is organized according to

the WordNet hierarchy), their study, known as Deep Expectation (DEX), took first place in the ChaLearn LAP 2015 competition on apparent age prediction [15]. Published the same year, [16] stated that due to a scarcity of photographs for the specified age range, the accuracy in age assessment of kid pictures was noticeably lower than that of its counterparts for other age groups. The use and distribution of raise serious legal and moral considerations.

2.0 LITERATURE REVIEW

In order to save time and money, it is crucial for the research community to share datasets. To train machine learning models and forecast additional information from incoming data, facial image datasets annotated with age and gender are required. [13] emphasizes the value of dataset sharing within the research community in favor of reproducing findings.

Datasets from earlier studies on age estimation have been made available. [14] acknowledged the lack of facial data and provided a dataset that gathered real-world images from open-source Flickr albums to address this shortcoming. Eight age groups were specified after the age and gender were annotated through observation.

In order to create a dataset of images annotated with actual age and gender, crawled more than 500,000 celebrity images from Wikimedia and IMDB in 2016. With a convolutional neural network (CNN) with 16 weight layers that was pre-trained on ImageNet (an image dataset that is organized according to the WordNet hierarchy), their study, known as Deep Expectation (DEX), took first place in the ChaLearn LAP 2015 challenge on apparent age prediction [15]. Published the same year, [16] acknowledged that due to a lack of images for the specified age group, the accuracy in age estimation of child pictures was noticeably lower than that of its counterparts for other age groups. The use and distribution of raise serious legal and moral issues.

2.1. DIGITAL FORENSIC BACKLOG

In order to create a collection of photos tagged with true age and gender, [9] crawled more than 500,000 celebrity photographs from Wikimedia and IMDB in 2016. With a convolutional neural network (CNN) with 16 weight layers that was pre-trained on ImageNet (an image dataset that is organized according to the WordNet hierarchy), their study, known as Deep Expectation (DEX), took first place in the ChaLearn LAP 2015 competition on apparent age prediction [15]. Published the same year, [16] stated that due to a scarcity of photographs for the specified age range, the accuracy in age assessment of kid pictures was noticeably lower than that of its counterparts for other age groups. The use and distribution of raise serious legal and moral considerations.

2.2 HUMAN FACIAL AGE PERCEPTION

The age of other people can be predicted rather accurately by humans. According to measurements, the error rate might range between 2.07 and 8.62 years depending on a number of variables, including the age of the assessor and the patient as well as the age gap between the two [21]. Young people's ages are frequently exaggerated [22], [23], [24], and there may be a desire to compare one's actual age to the projected age [25], [26]. An investigation of the own-age bias in the correct estimate of faces was provided by Moyse and Bre'dart [21]. The 114 participants' estimates of the ages of people in their own age groups (10–14, 20–30, and 65–75 years) were judged to be more accurate by the authors.

2.3 AGE ESTIMATION

In an era where processing these cases consumes a large portion of the time of digital investigators, it is crucial to identify the victim's age in a CAM subject. Due to differences in facial and physical characteristics, the lack of trustworthy cues, inherent ethnic differences, and the environment in which the victim appears, age as a soft biometric trait is difficult to predict [28]. The aforementioned study considers a number of variables that can determine whether or not an image is considered to be indecent as well as the appropriate age group. There have been countless studies done on age estimation. We have taken into

account the Mean Absolute Error (MAE), which is the discrepancy between the estimated age and the actual age, to gauge the accuracy of age estimation. These error rates have significantly decreased over the last 20 years. The published MAE rates have fluctuated between 1 and 5 years since early 2002. In 2014, [18] achieved a MAE of 1.47 by fusing a number of cutting-edge classifiers, including Fisher's LDA, Neural Networks, SVM, and others, using AdaBoost. However, because this study was conducted using a small, private dataset of 50 female images with ages ranging from 10 to 19, it highlights the dearth of information that is available as a result of the limited number of images of children and young adults that are available. and the moral ramifications of using their images. The average MAE for other algorithms that used the FG-NET dataset in 2011 was 4.1, which was achieved by [29] in that year. The used Contourlet appearance model was quicker and more precise at locating facial landmarks. as opposed to Active Appearance Models. [30] acknowledged the lack of accuracy when attempting to age-estimate children's faces through human observation. Additionally, male age predictions were more accurate after the age of 11 and female age predictions were more accurate in younger age groups. The claims, though, are based on a small sample that was kept in a secret dataset.

2.4 TRANSFER LEARNING

To avoid the time-consuming task of training data and minimizing the cost of running algorithms on hardware, numerous researchers have published pre-trained models. We can use pre-trained models created by other researchers thanks to the new learning framework known as transfer learning. Due to the lack of age-labeled face images, [31] took advantage of the transfer learning strategy to train deep convolutional neural networks. They claim that pre-train and fine tune are two steps in the transfer learning process, the former of which involves training randomly initialized networks with a sizable amount of labeled data and the latter of which involves using the learned parameters from the first process as an initialization for a new task. Pre-trained models are simply models developed to address a particular issue, and they are vulnerable to The Caffe Model Zoo has a variety of age estimation models that have been thoroughly documented and pre-trained. In 2015, [32] published a deep convolutional neural network for determining an individual's age and gender. To improve performance when learning representations from sparse data, their model was trained using the Flickr dataset of real-world facial images [14]. Using a ROBUST GPU, training each network took about 4 hours. Similar to [14], [33] also used the Audience dataset to train a ranking CNN-based framework for age estimation. Finally, we consider a pre-trained model that is compatible with the Caffe framework but not part of the model zoo. With a CNN, the Deep Expectation (DEX) model attempts to automate the estimation of facial age [9].

2.5 FACIAL AGE DATASET

The various algorithms combined in their work have been assessed using the open-source FG-NET dataset in a recent study by [34]. Furthermore, they claim that this dataset has a bias in favor of young children. This fact serves as inspiration for our research. However, the figures provided by FG-NET for images of minors fell short of what we had anticipated. The MORPH dataset includes about 78,000 pictures of people who range in age from 15 to 77. This dataset is relevant to our work when we consider age estimation for the teenager faction and is useful for facial recognition. As opposed to large datasets where feature tagging is uncommon, age and gender labels are well documented. Similarly, MEDS [11] is a dataset of mugshots of people who have passed away that includes the oldness feature. The various algorithms combined in their work have been assessed using the open-source FG-NET dataset in a recent study by [34]. Furthermore, they claim that this dataset has a bias in favor of young children. Although this is a Moti annotated, people under the legal drinking age should not read it. The FERET dataset, which is relevant for face detection [10], contains roughly 14,000 images. Age labeling is solely based on observation; as such, it is a validating fact for our study. Though, the Due to the high MAE values for age prediction produced by cutting-edge age estimation algorithms, research cannot rely on conjectures. As a result, we have thought about not using this dataset. The IMDB-WIKI dataset shared by is the largest dataset we are aware of. Given that the images were obtained by scraping images from the Internet, which is how this dataset's more than 500,000 celebrity photos were obtained, we have been cautious about using these photos because of copyright restrictions. We had noise in the source provided by DEX for our study, so we had to use a filter to get around this problem. Additionally, our solution decreased the number of images that were accessible by age cluster, and the absence of underage images was unavoidable.

A publicly accessible collection of labeled photos from Flickr's "in the wild" facial image database is called the OUI-Adience set. Despite [14] claiming to use Creative Commons license for their images, we have found that 89.55% of a sample of 10,842 images are associated with images that have copyright. We have chosen to omit these images, but we have acquired substitute images that are consistent in their tags and titles, can support the information above the child's age, and are covered by creative commons licenses.
The age dataset generator software was used to create a combined dataset from a number of constituent datasets (IMDB, WIKI, FG-NET, MEDS), and it was discovered that the volume of images was insufficient for both the young and the elderly. As an illustration, there are approximately 100 images available for each age range between 0 and 6 for males and 0 and 7 for females in the aforementioned combined dataset. Within a pool of more than 500,000 images, the program chooses distinctive images at random from various datasets. When a user filters noisy images, the values will typically decrease due to the software solution's continuous operation.

2.5.1 YAHOO FLICKR CREATIVE COMMONS 100M.

The Yahoo Flickr Creative Commons 100M (YFCC100M) was made available in 2014 [12] for the research community's use. This is the largest dataset of images and videos that has been made available for scientific research, as far as we are aware. The dataset is volatile due to the size of the collection, but taking into account the identification keys stored in the set, an updated response to a query to the Flickr API is possible. The sheer records were initially kept in a single text plain file. A script was run to manage the volume of data by iterating the repository line by line and copying each record to a NoSQL MongoDB collection. It was necessary to create pertinent indexes so that queries could be processed quickly. Once the database was established, it was possible to retrieve information from the collection, create URLS that allowed for public domain image access, filter images based on tags, and select the appropriate Creative Commons license. This dataset is useful for our study because it allows us to collect photos with creative commons licenses of people from underrepresented age groups.

3. EXISTING TOOLS AND MODELS

The effectiveness of age estimation is assessed using both online and offline tools. The main benefit of using cloud-based biometric services is that the results are processed by cutting-edge classifiers created by seasoned businesses in the field, like IBM, Microsoft, and Amazon. The primary drawback of online tools is the ongoing expense of using them. The majority of service responses are set up in JSON, which makes it simple for us to incorporate our performance evaluators.

In 2010, Amazon purchased "Rekognition" from Orbeus, a startup in artificial intelligence [35]. Artificial Neural Networks (ANNs) were used by the company to develop facial recognition software that identified characteristics in images. ANNs are systems that develop new skills by observing examples rather than by following a predetermined algorithm. They are organized into a final layer of output neurons, one or more hidden layers, and an initial input layer of neurons. To help non-experts train models without subject-matter knowledge, machine learning as a service was developed.

As a dual class output, the service's deep learning-based image analyzer can identify age with a minimum and maximum value. The tests were run assessing the mean average error with the minimum, maximum, and mean value of the output range after we evaluated the most suitable results and before normalizing the output to a single rate. Following our investigation, we decided to use the minimum value as a threshold for the procedures in a digital forensic case where the potential cost of inaccuracy is high.

As mentioned in Section 2.4, Deep Expectation (DEX) is a pre-trained Convolutional Neural Network model that outperformed the competition in terms of ratio. The main benefit of using offline pre-trained models is the flexibility to perform as much estimation as is necessary. The Chalearn LAP 2015 challenge was won by the proposed method by [9], which made use of the architecture of the aforementioned model [36]. Since the results are better than the mean value between the minimum and maximum age, the response must be normalized but differs from Recognition in this regard. The Support Vector Machine (SVM) algorithm is used by the free online service Kairos to separate various face types into the corresponding age

groups. In contrast to the other age estimation services, the performance is poor. Finally, a multi-layered deep learning methodology is used by Microsoft Azure Cognitive Service .

4. OVERVIEW OF THE SOFTWARE SOLUTION FOR PER- FORMANCE EVALUATION

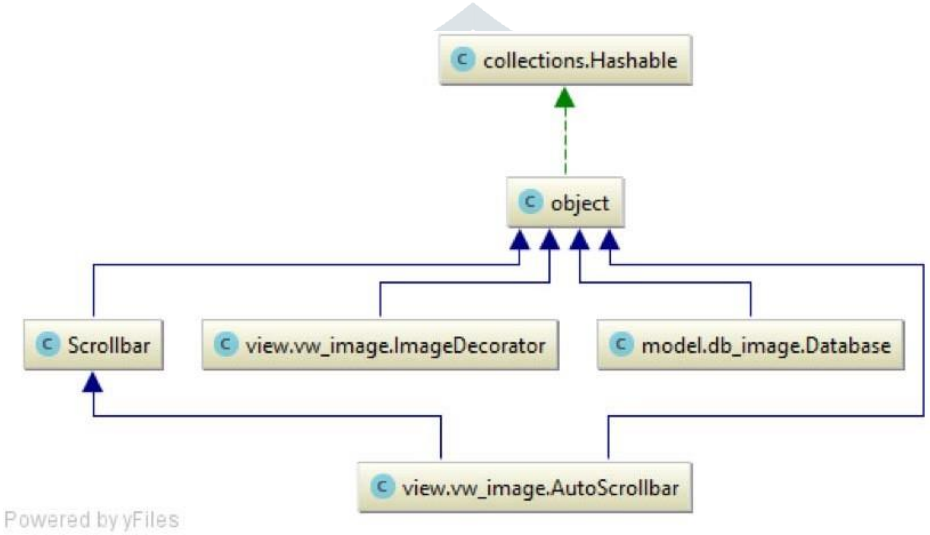


Figure 1:Decorator Pattern UML Class Diagram

A remarkably large number of labelled digital photos was necessary in order to assess the cutting-edge cloud-based biometric services. By choosing randomly distinct photographs from the various datasets specified in Section 2.5, a non-biased collection of images was created. Minimum and maximum ages, as well as the quantity of photos, are used to get results. To accommodate future adjustments to new datasets and services, the programmed needed to be scalable. As a result, the Model-View-Controller (MVC) software architecture was suggested, which allowed for code reuse and concurrent development. Scalability was allowed thanks to inheritance and many design approaches. See Figure 1 for an illustration of a UML class diagram using the decorator technique in our work. Amazon Rekognition2, Microsoft Azure Cognitive Services3, KAIROS4, and DEX 5 are the services used for performance evaluation. Future assessments can easily include new services thanks to inheritance.

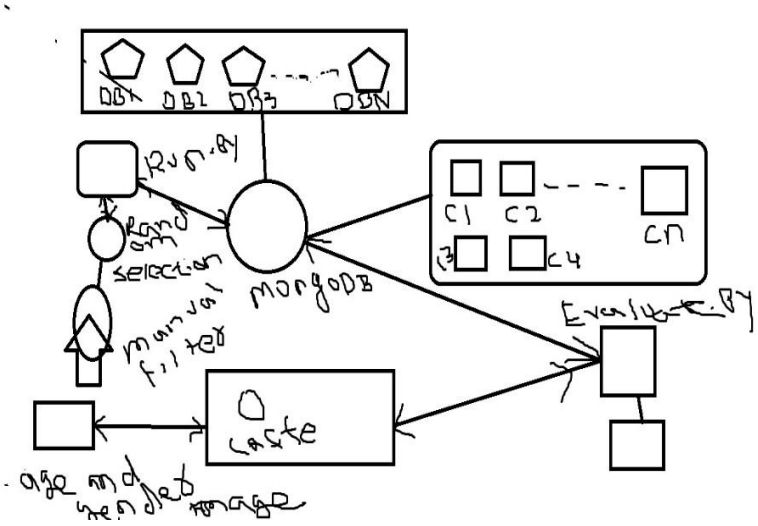


Figure 2: System Architecture

Multiple file sources (FG-NET, FERET, IMDB, WIKI, MEDS, YFCC100M, etc.) that correlate to each picture dataset make up the system architecture. To correspond with each file set, a distinct collection held the information. To help with the work, Python scripts were created, and in certain instances, Comma Separated Value (CSV) files were utilized to communicate with the NoSQL repository and the scripts. Both offline and cloud-based biometric services were assessed using an evenly distributed by age and gender collection of photos that was randomly chosen. The architecture of the system is shown in Figure 2.



Fig 3:Dataset Generator Software

The manual filtering phase is a crucial part of the system design. A dataset containing noise created by the combination of different random photos for each age group can only be efficiently filtered by human participation. As shown in Figure 3, users are given an interface where they may select each picture button to randomly create new images and delete any photos that are not relevant.

5. EVALUATION METHODOLOGY

Empirical assessment based on observation is the method of evaluation utilized in the study. The aim of the evaluation is to identify the pre-trained models and cloud-based biometric services with the lowest MAE. The evaluation's findings would help in deciding which is the most effective or show what mix of various services would help in developing a data fusion/ensemble strategy. Our study makes use of DEX's pre-trained Caffe model to predict the gender and age of digital photos. Comparative analyses were conducted using this tool and other cutting-edge online age predictors to assess the effectiveness and precision of the various forecasts.

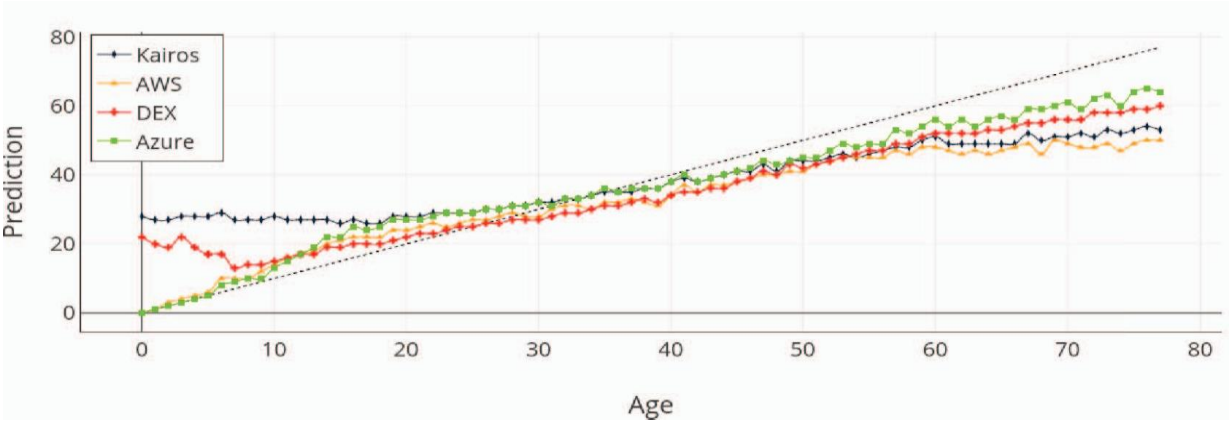


Figure 4. Average Estimated Age Compared with Actual Age across Entire Dataset.

In order to prevent imbalanced analysis, each of the cloud-based biometric services described for age estimation was assessed using a randomly distributed dataset. The same amount of photographs for each age group has to be utilised in order to ensure an equitable distribution. As a result, the dataset's overall size was constrained since not enough photos were available for some age groups. Initially, we meant to look at the entire age range from 0 to 100, however owing to a shortage of photographs, we only looked at the 0 to 77 age range. Additional photographs were manually gathered due to the dearth of images of children between the ages of 0 and 14 and the relevance of this age range to CAM research. licensed by the creative commons on Flickr, along with their age and gender. Based on the descriptions, the photo's title, the associated tags, or any other visual cues, photos were manually labelled with the subject's age and gender. In the end, this procedure made sure that 65 photographs could be included for each age and gender, yielding a dataset with 10,140 images overall.

The output from each of the four systems was recorded after each image had been processed. The outcomes of this evaluation under three influencing factors are reported in Section 6. These were then examined. The error rates that each system displayed, or the discrepancy between each subject's anticipated age and their actual age, were first compared throughout the whole age range. The information was then split up by gender to see if this had any impact on how accurately the age estimates were made. The dataset was separated into several age brackets for the third test (e.g., 0–9, 10–19, 20–29, etc.). To determine if particular systems performed better throughout a variety of ages, or whether one system Among the full dataset, might be regarded to be the most accurate.

6. RESULTS

The findings from our analysis of the available offline and online age estimation alternatives are presented in this section. First, in Section 6.1, we show the findings for the complete age range of our dataset (from 0 to 77 years old). The results of segmenting the dataset by gender are then shown in Section 6.2.

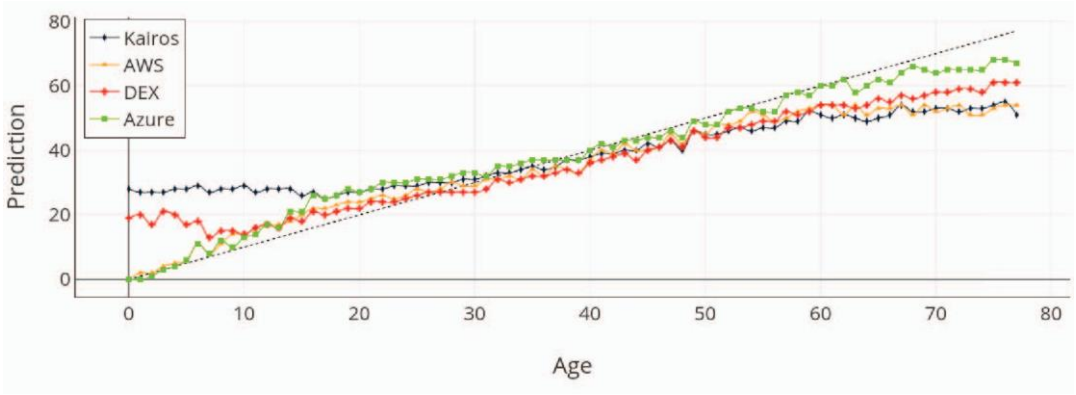
Finally, in Section 6.3, we contrast how well the four systems performed at various ages.

6.1 Entire Age Range Estimation

In order to determine which services are most useful, the first study measured the accuracy associated with each of the four systems throughout the full dataset. First, the MAE for each system across all participants was determined. Table 1 displays the outcomes of this.

Microsoft Azure was shown to produce the best outcomes with the lowest MAE using this criteria. This may be taken to suggest that there was an average discrepancy of 7.614 years between the anticipated age and the actual age throughout the complete range of all disciplines. The error rates for the other services were greater, with DEX outperforming AWS, which in turn outperformed Kairos, in terms of total performance. Although this is a helpful discovery in and of itself, a deeper look is needed to investigate the specifics of each service.

(a)



(b)

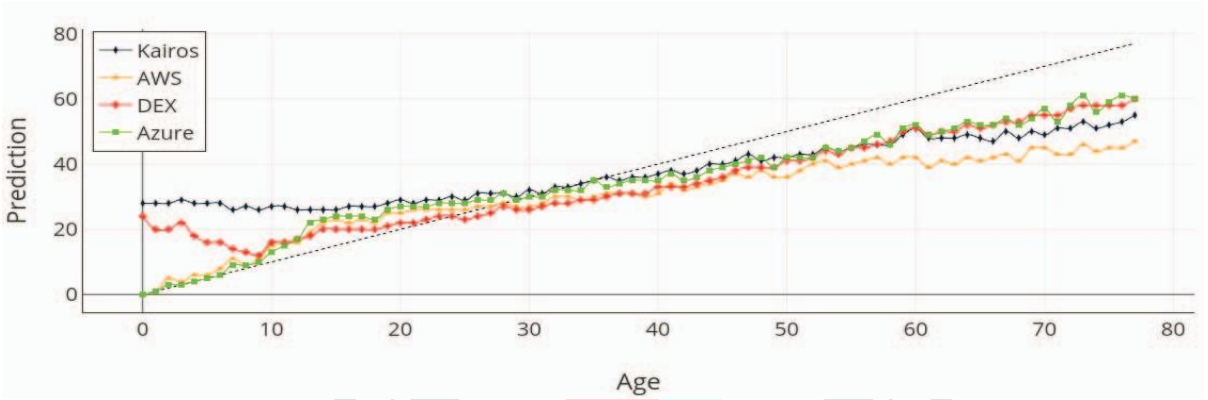


Figure 5. Average estimated age compared with actual age for (a) male subjects (b) female subjects.

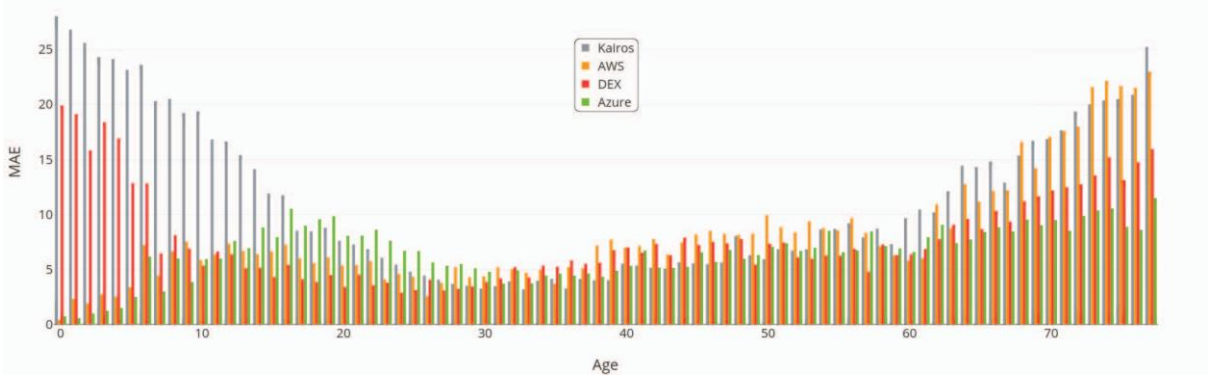
Figure 4 more clearly depicts how the four systems performed over the dataset. The participants' real ages are shown on the X-axis. For each service, a line is shown showing the average age it projected for subjects in each age group. Thus, each point reflects the expected average age for 65 participants of the same age. Correct predictions should fall along the dotted line, which is used to show this.
From this graph, a variety of intriguing conclusions may be drawn. Young children's ages are frequently substantially overestimated by DEX and Kairos. In the case of Kairos, this overestimation lasts until the late adolescent years, after which time its forecasts start to resemble those of its rivals.

Many interesting implications may be taken from this graph. DEX and Kairos commonly significantly overestimate the ages of young toddlers. This overestimation in Kairos' instance persists until the late adolescent years, at which point its predictions begin to approach those of its competitors.

Numerous insightful inferences may be made from this graph. The ages of young infants are commonly and significantly exaggerated by DEX and Kairos. For Kairos, this overestimation persists until the late adolescent years, after which its predictions begin to match those of its competitors.
It is useful to concentrate on the late adolescent years in particular, around the time when people cease to be minors in different countries, while considering digital forensic use-cases. Around age 10, the extremely accurate performance of AWS and Azure starts to deviate from the proper prediction line. The period between the middle and late adolescent years and early adulthood is when DEX, which does not work well on small children, is most similar to this ideal line. In Section 6.3, the relative performance of the systems over a variety of ages is further examined.

6.2 INfIUENCE OF GENDER ON ESTIMATION

(a)



(b)

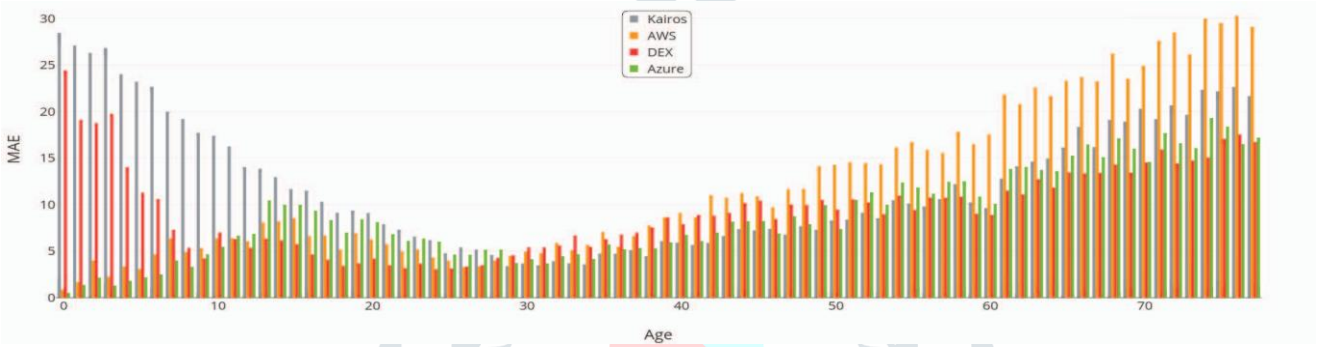


Figure 6. Mean Average Error Rate for (a) Male and (b) Female Subjects.

2. MEAN ABSOLUTE ERROR PER GENDER PER SERVICE.

Service	Male	Female
Kairos	10.6838	11.7960
AWS	7.2192	11.4057
DEX	7.1975	8.9613
Azure	6.4205	8.8092

We separated the dataset into equal gender groups and performed an analysis akin to that in the preceding section in order to further study the features of the four services. Table 2 displays the overall MAE for each service. The major intriguing finding from this table is that all four services consistently show a greater rate of mistake for female individuals than for male subjects. The only one of them, Kairos, has a gender disparity in mistake rates that is, on average, smaller than a year. The disparity is especially obvious for AWS, whose predictions for female individuals are more equivalent to Kairos but whose error rate for male subjects is just slightly higher than for DEX. the close friend

inquired past the general mistake rates. Similar to Figure 4, Figure 5 is built, with the exception that it displays distinct line graphs for each gender. Only the findings for male individuals are shown in Figure 5(a), and only female subjects are shown in Figure 5(b).

Both figures basically show the broad trends shown in the preceding section: Young respondents are overestimated by Kairos and DEX, while older people are often underestimated by all four services. The latter impact, however, manifests itself far more quickly in female individuals. By the age of 36, on average, all four services are underestimated, and as people become older, the difference gets worse. Azure, though,

is still far closer to the optimal line for masculine subjects.

6.3 Age Range Analysis

This is the main topic of the section that follows, where this data is seen from various age perspectives.
Before error rates start to fall in the 20s and 30s, a local peak is seen in the teenage years. Given the use cases that call for accuracy across this broad grownup age range, it appears that more attention will need to be paid to this area in the future.

6.4. In order to offer insights into the data, Figure 7 was created. Each service within each age group is displayed as a box in this graph. The average anticipated age for each actual age was used as the data for each box. The boxes display the median, interquartile range, and mean. Age Analysis by Range e, with the whiskers representing

The upper and lower bounds of the interquartile range being within 1.5 times. Where applicable, outliers are displayed as separate points.
This study's main goal is to identify the system or systems that accomplish the age prediction job with the greatest degree of accuracy. Even though Azure has been proved to have the lowest overall error, this number shows that this does not necessarily translate to a performance that is higher across all age groups. Table 3 provides an overview of the optimal strategy for each age group. Only for the youngest and oldest age groups is azure the most accurate. Unexpectedly, Kairos is the most accurate on average in the 30-39 and 40-49 age brackets, despite having the greatest total inaccuracy rate.



Figure 7. Mean Average Error Rate for Each System in Different Age Ranges.

7. CONCLUDING REMARK

There are several machine learning-based techniques aimed at reducing the backlog of digital evidence. The accuracy of automatic age estimate in photographic and video information must be improved as part of this endeavor. In this study, four distinct age prediction services—including cloud-based and offline sources—are evaluated. We may infer from the evaluation's results that gender is a soft biometric characteristic that has a substantial influence on the overall performance of the age prediction models since female individuals had a greater error rate. Male individuals' age estimations were generally more accurate and had an MAE that was around 2.1 years higher than those of the female participants.

It was crucial to understand how estimation services behaved across various age groups. Thus, three separate evaluations—individual age estimate, gender's impact on age estimation, and a number of grouped age ranges—were carried out. Our analysis demonstrates that, despite the Microsoft Azure online evaluation service dominating with the lowest overall error, the other alternatives performed better in a number of particular age categories.

This finding promotes more research in the field and shows that it may not be wise to depend just on one age estimating service. Researchers only have a small number of photos with gender and age labels that are precise. By integrating existing datasets, we presented a random, balanced dataset generator to get around this obstacle. Additionally, we have incorporated a library of digitised photos of minors, which has enabled us to objectively compare performance between boys and girls. This dataset will be made accessible for use in age prediction research and other areas, such as a resource that can be used to improve pre-trained algorithms.

7.1 Future Work

It is evident that picking the optimal face age estimate approach involves more consideration than just selecting the one with the lowest overall error rate, and that gender and age play a significant role in determining how well the four systems evaluated perform.
The fact that different systems perform well at different ages inspires more research into how the outcomes of many systems may be merged to increase the precision of forecasts as a whole. There are several machine learning regression approaches that can take the system predictions as inputs and ideally provide a forecast that is more accurate than the individual systems in terms of the individuals' actual ages.

Automating the laborious process of analyzing digital data from confiscated devices is the ultimate goal of this endeavor. Therefore, our goal is to research how automated machine learning-based solutions might help with digital forensic cases. Our goal is to broaden this study through the comparative comparison of more services as future work.
Furthermore, if the provided snapshot is not very clear, the shortcomings of the current instruments are demonstrated. due to the angle at which it was photographed and/or the inadequate illumination. We want to solve these common issues with facial recognition technology in the future.

REFERENCES

[1] B. Yang, J. Yan, Z. Lei, and S. Z. Li, “Aggregate channel eatures for multi-view face detection,” in IEEE International Joint Conference on Biometrics, 2014, pp. 1-8.
[2] P. Viola and M. J. Jones, “Robust real-time face detection. International journalofcomputer vision,” vol. 57, no. 2, pp. 137-154, 2004.
[3] M. T. Pham, Y. Gao, V. D. D. Hoang, and T. J. Cham, “Fast polygonal integration and its application in extending haar-like features to improve object detection,” in IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 942-949.
[4] Q. Zhu, M. C. Yeh, K. T. Cheng, and S. Avidan, “Fast human detection using a cascade of histograms of oriented gradients,” in IEEE Computer Conference on Computer Vision and Pattern Recognition, 2006, pp. 1491-1498.
[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in neural information processing systems, 2012, pp. 1097-1105.
[6] Y. Sun, Y. Chen, X. Wang, and X. Tang, “Deep learning face representation by joint identification-verification,” in Advances in Neural Information Processing Systems, 2014, pp. 1988-1996.
[7] S. Yang, P. Luo, C. C. Loy, and X. Tang, “From facial parts responses to face detection: A deep learning approach,” in IEEE International Conference on Computer Vision, 2015, pp. 3676-3684.
[8] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, “A convolutional neural network cascade for face detection,” in IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 5325-5334.
[9] C. Zhang, and Z. Zhang, “Improving multiview face detection with multi-task deep convolutional neural networks,” IEEE Winter Conference on Applications of Computer Vision, 2014, pp. 1036-1041.
[10] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, “Retinaface: Single-stage dense facelocalisation in the wild,” ArXiv, vol. abs/1905.00641, 2019
[11] Zhang K, Zhang Z, Li Z, et al. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks [J]. IEEE Signal Processing Letters, 2016, 23(10):1499-1503.
[12] Xiang J, Zhu G. [IEEE 2017 4th International Conference on Information Science and ControlEngineering (ICISCE) – Changsha (2017.7.21-2017.7.23)] 2017 4th International Conference on Information Science and Control Engineering

(ICISCE) - Joint Face Detection and FacialExpression Recognition with MTCNN[C]// International Conference on Information Science &Control Engineering. IEEE Computer Society, 2017:424-427

[13] B. Thomee, D. A. Shamma, G. Friedland, B. Elizalde, K. Ni,D. Poland, D. Borth, and L.-J. Li, "Yfcc100m: The new data inmultimedia research," Communications of the ACM, vol. 59, no. 2, pp. 64–73, 2016.

[14] C. Grajeda, F. Breitinger, and I. Baggili, "Availability of datasets for digital forensics–and what is missing," Digital Investigation, vol. 22, pp. S94–S105, 2017.

[15] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimationof unfiltered faces," IEEE Transactions on Information Forensics andSecurity, vol. 9, no. 12, pp. 2170–2179, 2014.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv pre-print arXiv:1409.1556, 2014.

[17] Ohol, M. R. M., & Ohol, M. S. R. PCA Algorithm for Human Face Recognition.

[18] Kasar, M. M., Bhattacharyya, D., & Kim, T. H. (2016). Face recognition using neural network: a review. International Journal of Security and Its Applications, 10(3), 81-100.

[19] Mikhaylov, D., Samoylov, A., Minin, P., & Egorov, A. (2014, November). Face Detection and Tracking from Image and Statistics Gathering. In Signal-Image Technology and InternetBased Systems (SITIS), 2014 Tenth International Conference on (pp. 37-42).

[20] Liu, Z., & Wang, Y. (2000). Face detection and tracking in video using dynamic programming. In Image Processing, 2000. Proceedings. 2000 International Conference on (Vol. 1, pp.

53-56). IEEE.

[21] IEEE ,Boda, R., & Priyadarsini, M. J. P. (2016). Face Detection and Tracking Using KLT And Viola Jones. ARPN Journal of Engineering and Applied Sciences, 11(23), 13472-1347

[22] E. Moyse and S. Brédart, "An own-age bias in age estimation of faces," Revue Européenne de Psychologie Appliquée/European Re- view of Applied Psychology, vol. 62, no. 1, pp. 3 – 7, 2012.

[23] J. B. Pittenger and R. E. Shaw, "Perception of relative and absolute age in facial photographs," Attention, Perception, & Psychophysics, vol. 18, no. 2, pp. 137–143, 1975

[24] R. Henss, "Perceiving age and attractiveness in facial photographs," Journal of Applied Social Psychology, vol. 21, no. 11, pp. 933–946, 1991.

[25] G. Willner and P. Rowe, "Alcohol servers' estimates of young peo- ple's ages," Drugs: education, prevention and policy, vol. 8, no. 4,pp. 375–383, 2001.

[26] M. C. Voelkle, N. C. Ebner, U. Lindenberger, and M. Riediger,"Let me guess how old you are: effects of age, gender, and facial expression on perceptions of age." Psychology and Aging, vol. 27,no. 2, p. 265, 2012.

[27] Kavia, M. Manjeet Kaur, (2016). "A Survey paper for Face Recognition Technologies". International Journal of Scientific and Research Publications, 6(7).

[28] Harguess, J., Aggarwal, J.K., A case for the average- half-face in 2D and 3D for facerecognition, IEEE Computer Society.

[29] J. A. Kloess, J. Woodhams, H. Whittle, T. Grant, and C. E. Hamilton-Giachritsis, "The challenges of identifying and classifying child sex- ual abuse material," Sexual Abuse, p. 1079063217724768, 2017.

[30] K. Luu, K. Seshadri, M. Savvides, T. D. Bui, and C. Y. Suen, "Con- tourlet appearance model for facial age estimation," in Biometrics (ijcb), 2011 international joint conference on. IEEE, 2011, pp. 1–8.

[31] E. Ferguson and C. Wilkinson, "Juvenile age estimation from facial images," Science & Justice, vol. 57, no. 1, pp. 58–62, 2017.

[32] Y. Dong, Y. Liu, and S. Lian, "Automatic age estimation based on deep learning algorithm," Neurocomputing, vol. 187, pp. 4–10, 2016.

[33] G. Levi and T. Hassner, "Age and gender classification using con- volutional neural networks," in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) workshops, June 2015.[34] S. Chen, C. Zhang, M. Dong, J. Le, and M. Rao, "Using ranking-cnn for age estimation," in The IEEE Conference on Computer Vision andPattern Recognition (CVPR), July 2017.

[35] H. Han, C. Otto, and A. K. Jain, "Age estimation from face images: Human vs. machine performance," in Biometrics (ICB), 2013 Inter- national Conference on. IEEE, 2013, pp. 1–8.

[36] K. Rajesh and K. Ramesh, "Artificial intelligence–fact or fiction,"Computing NaNo, 2012