

A review of age estimation research to evaluate its inclusion in automated child pornography detection

Lee MacLeod
David King
Euan Dempster

This is a pre-copyedited version of a contribution published in Intelligent computing: proceedings of the 2020 Computing conference, Volume 1, edited by Kohei Arai, Sypriya Kapoor and Rahul Bhatia, published by Springer.

The definitive authenticated version is available online via https://doi.org/10.1007/978-3-030-52249-0_38

A Review of Age Estimation Research to Evaluate its Inclusion in Automated Child Pornography Detection

Lee MacLeod¹, David King¹ and Euan Dempster¹

¹Abertay University, Bell Street, Dundee, DD1 1HG, UK
{1102974,d.king,e.dempster}@abertay.ac.uk

Abstract. The uses of artificial intelligence (AI) today seem limitless. It has helped organisations understand their customers more, provide them with better, more tailored services, and helped people with disabilities understand the world they previously could not. There are also many areas of current research for the use of AI. Aiding law-enforcement when they must analyse evidence of an indecent nature is one example where the use of AI, if successful, could enhance detection of indecent images and also reduce the workload and stress on the law enforcement staff employed in such activities. Working with indecent images of minors is particularly stressful. This paper reviews the current stage at which artificial intelligence finds itself when estimating a person's age. By reviewing its accuracy, it is possible to evaluate the feasibility of its inclusion in an artificial-intelligence-aided evidence analysis tool. With artificial intelligence currently capable of estimating a person's age to within a few years, its incorporation would most certainly allow photographs to be analysed and flagged if anyone is suspected of being underage.

Keywords: Age estimation, child pornography, pornography, digital forensics.

1 Introduction

The use of artificial intelligence (AI) is one that is growing at a tremendous pace. What started as a trickle a few decades ago has turned into a steady flow with almost daily advances. Today's AI is now able to handle many tasks such as running a building's heating and cooling [1]. This is in part due to the increased power of the computer systems that AI runs on. Within the past five years, the move from CPU to GPU hardware has resulted in increased accuracy, in conjunction with decreased training and processing times. This has allowed traditional solutions to work faster and be more accurate whilst making it possible to find answers to more complex problems.

AI has become the buzzword of the day amongst large corporations (Facebook, Microsoft, Google, etc.). Due to their respectively large infrastructures, these corporations are at the forefront of current research. This has allowed for some AI to be so well adapted to their situation, they are not only able to handle current tasks, but are so well trained they can work completely autonomously, with only the most critical problems being handed to humans for review.

One example of this would be the AI Google is using to control its data centre cooling [2]. In 2016, DeepMind (Google's AI division) developed an AI-powered recommendation system to improve the energy efficiency of Google's data centres. In recent times, this recommendation system has evolved to implement these recommendations autonomously.

To achieve this, every five minutes the AI polls the thousands of sensors distributed throughout the data centre that make up the cooling system. The AI then identifies which actions will minimise energy consumption, whilst remaining within defined safety constraints. Despite only being in place for a short space of time, Google estimates energy savings of around 30 percent, with further improvements expected.

Learning to Navigate in Cities Without a Map [3] is another DeepMind example. This time, using photographs taken as part of Google Street View, they developed an AI capable of navigating from A to B without the aid of a map. This is to simulate how humans navigate, by using landmarks or remembering certain streets.

To achieve successful navigation, their agent consists of three parts: A convolutional network used for processing the Street View images and extracting visual features; a locale-specific recurrent neural network used to memorise the environment; and finally, an interchangeable locale-invariant recurrent network that decides the agent's journey.

As the name suggests, this third component is interchangeable and unique to each city where the agent navigates. All of this allows the agent to remember navigational cues from previous cities and apply that knowledge to cities not previously visited. The ultimate goal of this research is to aid scientists in the understanding of biological approach navigation.

Currently, AI is used for a wide variety of tasks, anything from the previously noted heating and cooling of buildings to describing an image's content to the visually impaired [4]. A more specific area of focus is the estimation of a human's age-based upon facial details present in their photograph. The practical uses of this may seem niche at first glance, but it can be used in a wide variety of functions, including biometrics and human-computer interaction [5, 6], and may provide further applications in child pornography detection.

This author's research seeks to establish practical means by which AI may be used as a digital forensics tool for analysing evidence regarded as child pornography. To say whether or not a child is present within gathered visual evidence, it is important to identify the ages of the humans within said evidence. This paper establishes the current state of age estimation through a review of the literature and hypothesises a possible approach to age estimation for the author's research project.

1.1 Significance of Research

The primary objective of this project is to invent a new, robust method of analysing visual pornographic evidence within digital forensic investigations, which has both inconsequential false positives and low false-negative detection rates. The speed of estimation is a secondary objective of this research, although, given the current state of computer power, it is inevitable that the new method will be faster than a human analyst.

Such a method would allow analysts to focus their efforts on other relevant materials, such as e-mails and other content-heavy files, which AI would struggle to handle.

1.2 Child Pornography

Before one may design an automated tool for the detection of child pornography, one must first understand what the term ‘child pornography’ means. Whilst at first glance, the term may seem easily definable, there is indeed some ambiguity surrounding its definition. Focusing on the term ‘pornography’ one would be forgiven for considering it to simply mean ‘images of naked people’. However, this could be considered an oversimplification when one considers Naturism, for example. Under this hypothetical, simplified definition of pornography, this legitimate and largely appropriate activity would be considered pornography, when in fact it is considered as the cultural (and sometimes political) movement that both encourages and defends social nudity.

Pornography may instead be broadly described as “Printed or visual material containing the explicit description or display of sexual organs or activity, intended to stimulate sexual excitement” [Oxford English Dictionary]. Now, taking into consideration the legal definition of a child includes their inability to provide consent to sexual activities, child pornography must be considered as illegal [7]. It may further be described as a portrayal of sexual assault against a child. In Scotland, Under Section 52, 52A Civic Government (Scotland) Act 1982 [8], the mere possession of material depicting such situations is considered a criminal offence, with consequences including up to 10 years in prison and a fine not exceeding £5,000.

Understanding the definition reveals the problem associated with current methods as well as the problems with developing a working solution. To correctly classify child pornography, a system must not just take into consideration a child’s nakedness but grasp the very situation in which a child finds itself. This is because even in situations where a child is fully-clothed, they may be presenting themselves in a provocative way and as such be considered as victims of child exploitation. Considering this renders current methods of adult pornography detection unusable for child pornography detection.

2 Research Methodology

2.1 Hypothesis

The hypothesis of this investigation is “It is possible to automatically estimate a human’s age based on their photographic image”. To prove this hypothesis, a detailed review of research of the last 10 years has been undertaken.

2.2 Approach to Research

To obtain a significant body of research to discuss, a chronological list of significant works was first produced from the last 10 years. Selecting works that obtained results with a low mean average error (MAE < 5 years) narrowed this initial list further. This step produced research that shared closely grouped mean average errors. Therefore, it was important to select research, as well as producing low mean average errors, that explored interesting techniques to achieve those results.

Of course, this method produced a very short list of results. Whilst the works within this list are both contemporary and have low MAE, it does not necessarily include seminal works. However, for inclusion in a digital forensics' toolkit, research that produces a low MAE is the focus rather than a paper's distinct methodology.

Before an appropriate detection system may be built, it is first imperative that a means of age estimation be found. It was, therefore, important to analyse the research from a digital forensics' perspective. Whilst none of the research reviewed is concerned necessarily with digital forensics, age estimation may be considered important when seeking a working solution for child pornography detection. Selecting research that presented a low mean average was important because it allows possible future detection solutions to not only estimate a broad age range but also produce a narrow age range that is accurate to within a few years.

2.3 Proposed Research

In a system designed to locate child pornography, age estimation plays an important role. However, age estimation is only one of many techniques present within this author's research. There is also nudity detection, filename analysis and investigating the metadata of a file to make sure it is what it eludes to be.

To successfully detect media files that contain child sexual abuse, this author proposes a toolkit in which files are analysed both with the use of artificial intelligence and without. Filename analysis is one example where simple programming could be used to check the file's name against a dictionary of keywords. Each mode is classed as an input into a Weighted Naïve Bayesian Formula. A weight is applied to each mode to convey the level of importance that any individual mode should hold over another.

Every mode with its associated weight is used as inputs into the Bayesian formula, which ultimately assigns the probability of the input file containing child sexual abuse. To evaluate this approach, firstly one will use a hard drive containing pornographic material that has been deliberately placed. This way, it is possible to measure the toolkit's performance by measuring the number of files found versus the total number of known files. Secondly, the toolkit will be given to digital forensic analysts within law-enforcement to gather real-world data with regards to its accuracy, performance, and qualitatively, how easy it is to use.

2.4 Age Identification

Jain, Doss et al. [9] describe age classification as “soft biometrics”. Unlike traditional methods of biometrics, the soft variety identifies features such as age, gender, skin and eye colour. Unfortunately, whilst calculable elements of the human physiology exist, it has been expressed in the past that no such methods exist for reliably calculating human age [9]. Nevertheless, promising advancements within the past five years bring this into question.

As an example, consider the interesting work conducted by Eugene Bobrov et al., *PhotoAgeClock: deep learning algorithms for development of non-invasive visual biomarkers of aging* [10]. Deep neural networks were trained on 8,414 high-resolution images of eye corneas labelled with their corresponding age. For people 20-80 years old, their model was able to achieve a mean absolute error of 2.3 years and 95% Pearson and Spearman correlation. This is just one work in many of recent years that continue to prove that AI age estimation is both practical and accurate.

Humorously, one recent social network-based trend sparked an interesting debate, with its genesis coming from a flippant post by Kate O’Neill of Wired magazine [11, 12]. The trend involves showing two pictures, side by side – the left, oneself 10 years previous; the right, oneself today. Whilst the general user regards this as harmless fun – a way of remembering past events – the debate has ignited around whether or not this is, in fact, a way for social media companies to have access to perfectly labelled pictures showing a wide variety of age progression over a fixed period. Whilst Facebook has been quick to refute these claims with a press release stating they have no part to play in the game’s orchestration, it does make for an interesting topic of debate when very few full and clean datasets currently exist.

To have a dataset that is both complete and accurate (such as having a person’s age correct) is important for training and testing AI because, without a dataset’s accuracy, the algorithm is sure to produce inaccurate results. Whilst such datasets do exist, using them to train an AI to estimate correctly a child’s age does present a problem because currently available datasets show a bias towards adult ages. To estimate a child’s age, however, there may be several possibilities. One such example would be to train an AI using an adult orientated dataset. Negative results produced by the AI would, therefore, would allow the human observer to infer the person is a child.

Kwon, Y.H., and da Vitoria Lobo, N.

This fascinating research paper produced in 1999 [13] details how the use of various ratios produced by measuring facial features can be used to estimate the classification group (baby, adult, or senior) a person belongs to (see Fig. 1). Whilst this paper does not use AI as its classification method, it is one of the first papers to discuss a mathematical implementation for age estimation.

To achieve an estimation, a several facial features are evaluated. Initially, the outline of the face is found by drawing an oval slightly below-centre of the middle (to compensate for hair). This oval shape is iteratively expanded by increasing the axes at half-lengths and updating the centre.

Using similar methods, the chin is located and the initial outline is adjusted to suit. With the outline of the face determined, it is now possible to find the sides of the face before finally finding the eyes, mouth, and nose. Once all locations are found, several ratios are formed from their locations. These ratios, in conjunction with wrinkle geography, are used to assign a person to an age group.

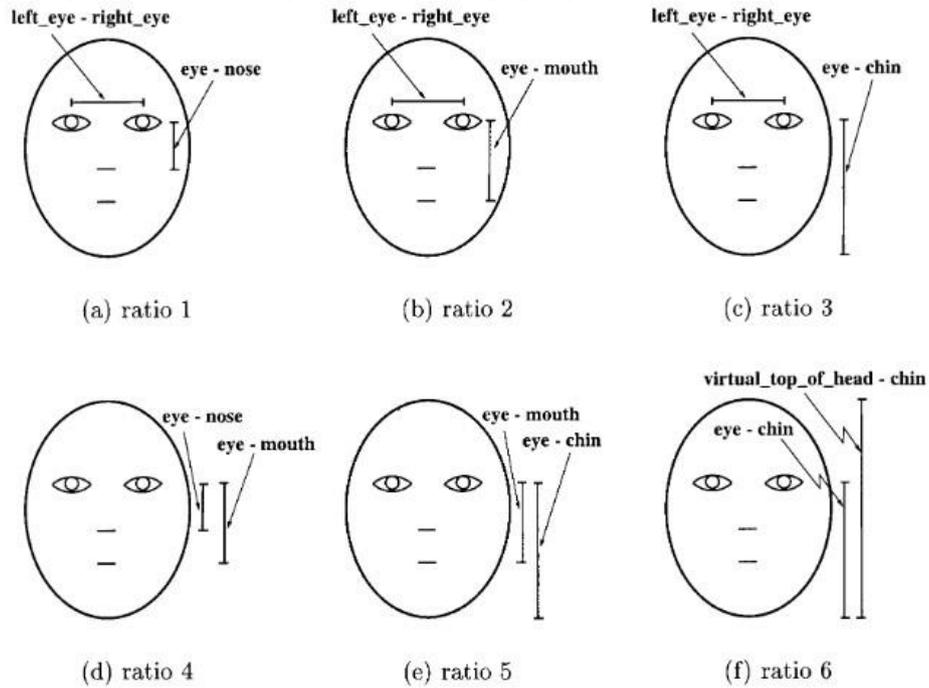


Fig. 1. The different ratios used. Each ratio measures different elements of a person's face and all together, along with wrinkle geography, are used to estimate a person's age.

The 'wrinkle geography' is important, as there is a direct correlation between a person's age and the degree of wrinkles. With the positions of the eyes, mouth, and nose now known, it is possible to map the remaining areas of the face as possible wrinkle locations. Wrinkle detection is achieved by using *snakelets* [14] (a snake segment designed to find curved segments of an image) that are dropped in random orientation and positions in the mapped areas. Any snakelets not found to be near wrinkles are removed. At this stage, wrinkles are marked by a sudden change in colour and depth. To determine a wrinkle's legitimacy, remaining snakelets are investigated by summing distances between each snakelet and wrinkle point. Sums found to be within a threshold distance are an indication of a wrinkle. These steps are repeated for all the remaining snakelet points until complete wrinkle geography is produced; with snakelets accurately following the path of the person's wrinkles. Shown in Fig. 2, is a systematic example of this process.

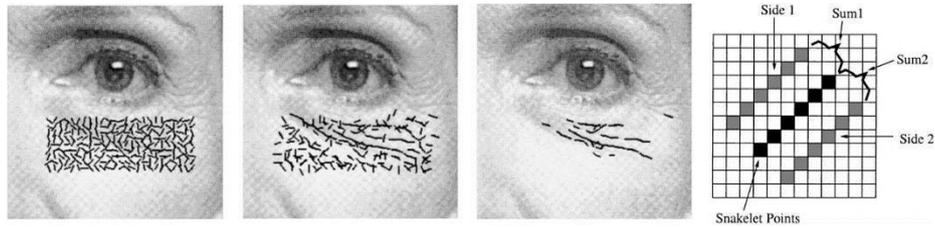


Fig. 2. Three states of wrinkle detection and the mapping process. The furthest left view shows the snakelet's initial state, which is being dropped in random orientations. The middle views show the snakelets after stabilisation. The final view demonstrates the summing process that is used to complete the wrinkle geography.

With the level of wrinkles for the test subject now determined, in conjunction with the other feature information, it is possible to determine a group in which the person belongs. To validate this approach, 15 faces (five for each baby, adult, and senior groups) were used for testing. For the 15 faces, the results show that “classifications were 100% correct.”

Whilst certainly interesting (and possibly questionable because of its high accuracy level), it is worth noting this was a simple categorisation problem. Very few wrinkles indicated the person was a baby, not a person with good skin quality. Using this approach to define an age sub-category (say 35-35, for example) for individuals would prove to be less accurate. Besides, the photographs used for training and evaluation were all ‘mugshots’ in which the face was square to the camera and completely visible. Such methods, therefore, would again yield poorer results when used for images that failed to match the ones used for this research.

Huerta, Ivan et al.

According to Huerta, Ivan et al. [15], existing works tackle age estimation by relying on complex, finely tuned models, which are reliant on several facial landmarks for increased accuracy. Instead, their age estimation approach uses points around the eyes. This is achieved by first successfully detecting the facial region within photographs. Using a convolutional neural network, they detect eyes and surrounding areas. Further work is undertaken to yield an aligned version of the face by a “non-reflective similarity image transformation”.

The choice of visual features that are extracted and subsequently used for classification directly affect the performance of the age estimation. For this research, the authors chose to select commonly used invariant descriptors, including Histograms of Oriented Gradients (HOG), Local Binary Patterns (LBP), and Speeded-Up Robust Features (SURF).

HOG [16], in particular, has been used with success in many computer vision applications. Simply, it works by dividing the image region into a grid. Each cell is further split into smaller cells. For each pixel within a cell, horizontal and vertical gradients are obtained. For the histogram, Q bins for the angle are chosen. This may be thought of as an array of angles, where unsigned orientation is usually used, such that

angles less than 0° are increased by 180° . The histograms are then concatenated, providing the final descriptor. At the end of this process, one is left with a complete and largely accurate outline of an object (Fig. 3). In Huerta's research, this results in an outline around a human head, bringing it into focus.

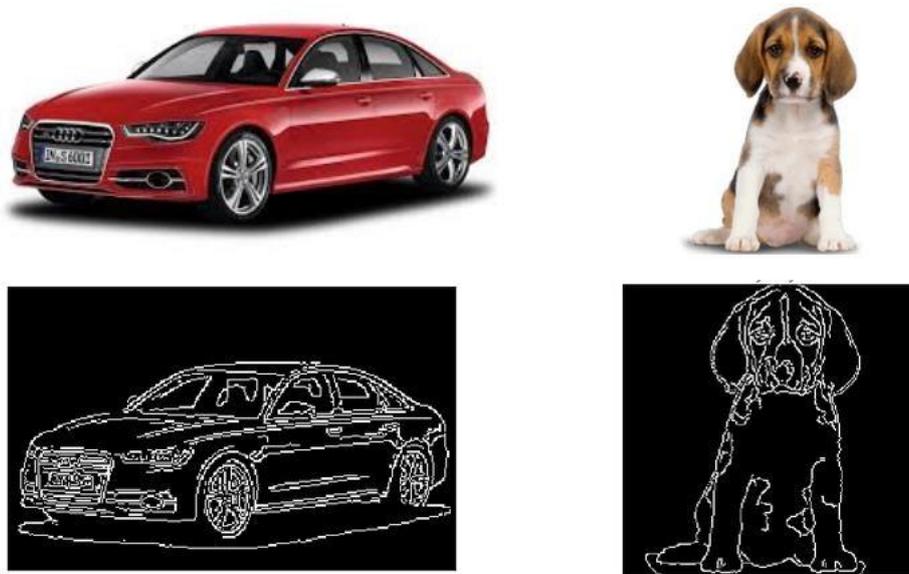


Fig. 3. Results of a sample HOG routine. The top row is the inputs to the HOG process and the bottom row shows their respective outputs. After the HOG process, one is left with a detailed outline of the input image. ¹

Local Binary Patterns (LBP), are a texture descriptor made popular by Ojala et al. in their 2002 paper *Multiresolution Grayscale and Rotation Invariant Texture Classification with Local Binary Patterns* [17]. LBPs compute a local representation of texture by comparing each pixel with its surrounding neighbourhood of pixels. This is unlike Haralick texture features that compute a global representation of texture.

To construct an LBP texture descriptor, the input image is first converted to grayscale. For each pixel in the image, a neighbourhood of r size surrounding the centre pixel is selected. If the intensity of the centre pixel is greater than or equal to its neighbour, the neighbour's value is set to one; otherwise, zero. The value for the centre pixel is calculated by first converting all of the neighbour's binary values to decimal before taking their sum. The output of this sum is used as the value of the centre pixel. This process is repeated until the complete image is analysed. Finally, to build the LBP descriptor, a histogram is computed over the result, in which each bin corresponds to an LBP code.

¹ www.analyticsvidhya.com/blog/2019/09/feature-engineering-images-introduction-hog-featuredescrptor/

Another method, Speeded-Up Robust Features (SURF) [18], as the name suggests, is a fast and robust method for image matching and object recognition, and a faster alternative to SIFT (Scale-Invariant Feature Transform).

In the case of Huerta, Ivan et al. they use U-SURF, an upright version of the technique. U-SURF is not invariant to image rotation and therefore performs better with images where the camera is unlikely to shift. The square image region to describe is partitioned into 4x4 sub-regions. Horizontal and vertical wavelet responses d_x and d_y are computed and weighted with Gaussian. The sum of these responses and their absolute values are stored, generating a four-dimensional vector

$(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$ for each sub-region, and these are concatenated to form the final 64-dimensional descriptor of the image region, SURF₆₄.

One complaint shared in many age estimation types of research is the limited freely available datasets. For this research, the authors focused on the MORPH dataset [19], given its large number of samples and age distributions. Whilst another dataset, FRGC [20], does exist however, it is noted that only one previous publication has used it. With the dataset chosen, training and evaluation could occur. To evaluate the accuracy of age estimators, Mean Average Error (MAE) and Cumulative Score (CS), both well-documented metrics, were used.

The results put forward for this relatively novel approach are promising too, with a MAE score of 4.25 years, compared to the previously published 4.38. The authors note that this could be further improved (to MAE 4.16) when using one descriptor and large images.

Liu, Hao et al. (2017)

Liu, Hao et al. 2017 paper [21] proposes a “group-aware deep feature learning” approach to facial age estimation. Compared to other approaches, which rely on handcrafted descriptors, this approach first split ages into 10-year segments. These age ranges are further split into smaller, overlapping ones (Fig. 4).

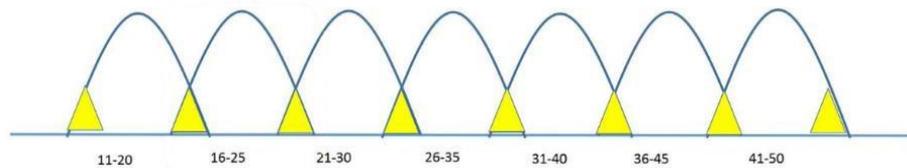


Fig. 4. An age group is split into three regions: one non-overlapping region in the middle and two symmetrical overlapping regions adjacent with neighbouring age groups coloured in yellow. The yellow regions are overlaps of four years.

This approach is taken because it would be difficult and impractical to collect a large enough dataset of facial data for the same person covering a wide range of age labels. Instead, as it is easier to collect facial images of individuals in short-intervals (e.g.

intervals of less than 10 years), their model splits age ranges into groups. The authors posit that the appearance of a person of 50 years old is more similar to those 47-52 years old than those 30 years old and therefore it is better to split their age process into several discrete groups.

For training, each image was resized to 244x244 pixels and then downsampled to 64x64 and 32x32 as multi-scale input to the defined multi-path convolutional neural network (CNN). To unify and align the images, all facial bounding boxes and facial landmarks – the central region of the eye, and the base of the nose – were detected. An overview is shown in Fig. 5.

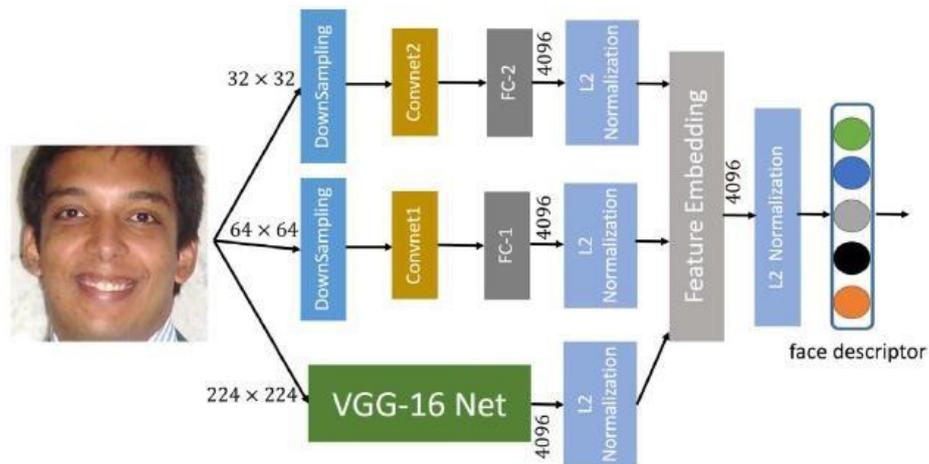


Fig. 5. The multi-path network architecture. The input image is scaled to 32x32, 64x64, 224x224 respectively. Each scaled version is passed through its respective subnet and normalised through the embedding layer. 4096 denotes the output dimension of each subnet. The face descriptor is the output layer, which contains the estimated age.

Initial training and testing took place on the FG-NET dataset [22]. This dataset contains just over 1,000 images from 82 persons, with ages ranging from zero to 69. Testing was conducted using a “leave-one-person-out” approach, such that one person’s dataset was omitted at random, with the remaining ones being used for models training. This was performed 82 times for cross-validation. The results were then averaged to obtain final age estimations. Using this data, the MAE was 3.93.

Further training and testing were conducted using the MORPH dataset. Containing some 55,608 images from about 13,000 subjects, and with ages from 16 to 77, it is significantly larger than the FG-NET dataset. Testing was undertaken slightly differently from previous, due to the dataset’s size. Instead, the dataset was split equally into 10 folds. One of the folds was selected at random and used for testing, whilst the rest were used for training. This process was repeated 10 times and the resulting average was taken. This time, the MAE was 3.25.

Finally, the ChaLearn Challenge dataset was used. This dataset contains just over 4,000 images and 1,500 images for validation. Unlike the previous datasets, all the images are taken in a large variety of settings, with further variance introduced by different aspect ratios and quality changes. When compared to the other datasets, the authors believe that the diverseness of the ChaLearn Challenge dataset causes the MAE to be slightly higher, 4.21. Regardless of the dataset used, Hao et al. were able to produce industry-average results at worst, with marginally more accurate results in most cases.

Hu, Zhenzhen et al. (2017)

Hu, Zhenzhen et al. present state-of-the-art results in their paper *Facial Age Estimation with Age Difference* [23]. Whilst typical approaches towards age estimation are centred on using labelled data, their approach uses age *difference* information instead. To obtain access to this type of information, the authors had to create their dataset. This was achieved by extracting photographs with their related information from Flickr.com. Importantly, this related information contains the dates the images were taken and uploaded. After filtering the images through face detection and alignment algorithms, the remaining 150,000 images create the final dataset.

A network is trained with the FG-NET and MORPH datasets to create a pre-trained age estimator. For their non-age-labelled dataset, two images of the same person are combined as a pair. For the age-labelled datasets, the authors follow the work undertaken by Geng et al. [24] and Geng [25]. This overarching approach was undertaken, as it should allow for consistent age estimation. For their non-age-labelled dataset, they used the year difference between two images of the same person to train an age difference estimator. The two images are assigned a probability of their age with a range of estimated ages. By using several loss functions, they can evaluate the age probability distributions and adjust the loss function for the network as a whole (Fig. 6).

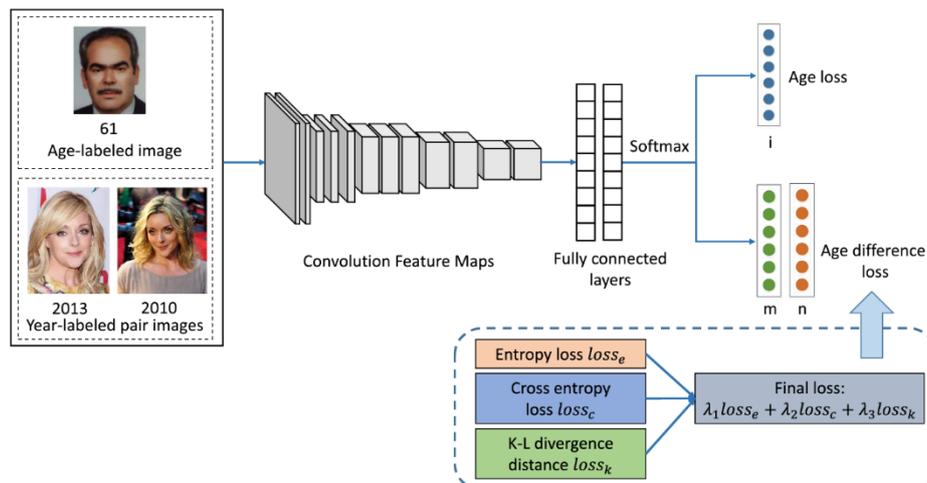


Fig. 6. Model displaying the input-output flow of the network. The model is pre-trained on the FG-NET and MORPH datasets. For the non-age-labelled dataset, two images of the same subject are combined as a pair. The network is then fine-tuned using the image pairs.

Much like other experiments, MAE and cumulative score (CS) are used as performance metrics. To evaluate the network’s performance, it was first tested using the age-labelled dataset. 80 percent of images were randomly selected as the training data, with the remaining images set aside for testing. MAE was 2.8 and 2.78 (FG-NET and MORPH, respectively), which are state-of-the-art (Table 1).

Further evaluation was undertaken using the year-labelled dataset. As this dataset does not contain ages, but rather years, it was only possible to test the age difference estimator. Like the previous testing, the dataset was split, with 30,000 images from 700 subjects selected. The MAE of the age difference estimator is 1.74 while the average age difference of the year-labelled dataset is seven years. To evaluate the usefulness of age difference information, results with and without this information were investigated. Without this information, the MAE was 3.12; whilst with this information, it dropped to 2.78 (Table 2).

Table 1. MAE Comparison with age estimation algorithms on the age-labelled databases.

Heading level	FG-NET	MORPH
RED + SVM [26]	6.33	6.57
SVR	5.45	5.52
kNN	8.13	8.22
CS-LBFL [27]	4.36	4.37
Our methods	2.8	2.78

Table 2. MAE comparison with age estimation algorithms on the year-labelled databases.

Methods	MORPH
GoogleLeNet with year-label data	2.78
GoogleLeNet (Gaussian label)	3.13
GoogleLeNet (Single label)	3.15
Huerta [28]	3.88
Wang [29]	3.81
Wang [30]	4.77

2.5 Inclusion in Further Research

As one can see, this area of research is very active. As such, there are many different methods of approaching the problem of age estimation. All methods discussed rely on common datasets such as FG-NET or MORPH for training and testing. Even Hu, Zhenzhen et al., still rely on these datasets despite constructing their own to reduce their

MAE. The different strides made in current research is interesting, considering their differences in accuracy, despite sharing these datasets.

Although the research covered discussed age estimation as a means of soft-biometrics, age estimation may also play an important role in digital forensics. Its inclusion in an automated child pornography detection tool is certainly worth considering as it would most likely aid in its successful detection. Furthermore, this could be achieved by implementing some of the methods discussed. For example, Huerta, Ivan et al. pose a novel approach, which, whilst indeed a simpler method when compared to the other research discussed, would provide reasonably accurate results when included as a single-mode in a multi-modal system. Furthermore, if a version of Hu, Henzhen et al.'s network could be successfully trained and tested, then-current estimation accuracy could be further improved.

In contrast, in situations where age estimation data containing children is not abundantly available, it may be possible to use a person's height instead. Within the past two decades, research has been conducted where estimating the height of a human has been the focus [31-33]. As a prerequisite for the vast majority of this research, however, data presented to any algorithm must include the known measurements of an object, often thought of as a pseudo-calibration. By doing this, algorithms can make presumptive estimations on objects whose dimensions are unknown by essentially comparatively scaling (or, indeed, downscaling) them. Consequently, for digital forensics, absolute measurements would most likely be impossible to obtain because they would lack the means of calibration in the field; as this calibration would have to happen at the time the photographs were being taken.

3 Conclusions

Using AI to estimate a person's age is a growing field, especially when it may be used as a by-product of other uses, such as facial recognition. The MAE has consistently been reduced over recent years, currently sitting at just under three. A decade ago, that MAE was as high as eight or nine. This low error rate, and inversely high accuracy rate, most certainly helps when one is to consider using age estimation as part of a child pornography detection toolkit.

As mentioned, however, the distinct lack of child-focused training data compounds the difficulty of accurate age estimation for younger people. Nevertheless, it may still be possible to estimate a child's age despite this. By detecting all persons within a photograph and estimating their respective ages, it would be a safe assumption to make that any persons below a certain age (e.g. 18) would be classed as a child. To detect a child within pornography, this approach, in tandem with other identifiers (such as their height), may yield a high enough accuracy where it would be possible to consider its inclusion in real-world applications.

3.1 Future Work

The goal of the author's research is to incorporate successfully AI into digital forensics, through the development of a toolkit. AI techniques discussed in this review paper show contemporary methods can estimate a person's age to within three years. This current accuracy means that age estimation must be included in the author's toolkit

To that end, the author will be including AI-based age estimation within the toolkit under development. Future work will include further research into suitable AI architectures and, once a suitable architecture has been selected, its inclusion into the toolkit.

References

- [1] Google is Switching to a Self-Driving Data Center Management System [Internet]. Data Center Knowledge. 2018 [cited 2019 Oct 31]. Available from: <https://www.datacenterknowledge.com/google-alphabet/google-switching-selfdriving-data-center-management-system>
- [2] Safety-first AI for autonomous data centre cooling and industrial control [Internet]. Deepmind. [cited 2019 Oct 31]. Available from: </blog/article/safety-first-ai-autonomous-data-centre-cooling-and-industrial-control>
- [3] Mirowski P, Grimes MK, Malinowski M, Hermann KM, Anderson K, Teplyashin D, et al. Learning to Navigate in Cities Without a Map. arXiv:180400168 [cs] [Internet]. 2019 Jan 9 [cited 2019 Oct 31]; Available from: <http://arxiv.org/abs/1804.00168>
- [4] Newton C. Facebook begins using artificial intelligence to describe photos to blind users [Internet]. The Verge. 2016 [cited 2019 Oct 31]. Available from: <https://www.theverge.com/2016/4/5/11364914/facebook-automatic-alt-tags-blindvisually-impaired>
- [5] Geng X, Zhou Z-H, Smith-Miles K. Automatic Age Estimation Based on Facial Aging Patterns. *IEEE Trans Pattern Anal Mach Intell.* 2007 Dec;29(12):2234–40.
- [6] Ran He, Wei-Shi Zheng, Bao-Gang Hu. Maximum Correntropy Criterion for Robust Face Recognition. *IEEE Trans Pattern Anal Mach Intell.* 2011 Aug;33(8):1561–76.
- [7] Taylor M, Quayle E. *Child pornography: an Internet crime.* Hove, East Sussex; New York: Brunner-Routledge; 2003. 236 p.
- [8] Civic Government (Scotland) Act 1982 [Internet]. [cited 2019 Oct 31]. Available from: <http://www.legislation.gov.uk/ukpga/1982/45/section/52>
- [9] Jain AK, Doss SC, Nandakumar K. Soft Biometric Traits for Personal Recognition Systems. In: Zhang D, Jain AK, editors. *Biometric Authentication* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2004 [cited 2019 Oct 31]. p. 731–8.

Available from: http://link.springer.com/10.1007/978-3-540-25948-0_99

- [10] Bobrov E, Georgievskaya A, Kiselev K, Sevastopolsky A, Zhavoronkov A, Gurov S, et al. PhotoAgeClock: deep learning algorithms for development of non-invasive visual biomarkers of aging. *aging*. 2018 Nov 9;10(11):3249–59.
- [11] Kate O’Neill on Twitter: ‘Me 10 years ago: probably would have played along with the profile picture aging meme going around on Facebook and Instagram Me now: ponders how all this data could be mined to train facial recognition algorithms on age progression and age recognition’ [Internet]. [cited 2019 Oct 31]. Available from: <https://twitter.com/kateo/status/1084199700427927553>
- [12] Facebook’s ‘10 Year Challenge’ Is Just a Harmless Meme—Right? | WIRED [Internet]. [cited 2019 Oct 31]. Available from: <https://www.wired.com/story/facebook-10-year-meme-challenge/>
- [13] Kwon YH, Lobo N da V. Age Classification from Facial Images. *Computer Vision and Image Understanding*. 1999 Apr;74(1):1–21.
- [14] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. *Int J Comput Vision*. 1988 Jan 1;1(4):321–31.
- [15] Huerta I, Fernández C, Segura C, Hernando J, Prati A. A deep analysis on age estimation. *Pattern Recognition Letters*. 2015 Dec; 68:239–49.
- [16] Histogram of Oriented Gradients | Learn OpenCV [Internet]. [cited 2019 Oct 31]. Available from: <https://www.learnopencv.com/histogram-of-oriented-gradients/>
- [17] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Machine Intell*. 2002 Jul;24(7):971–87.
- [18] Bay H, Tuytelaars T, Van Gool L. SURF: Speeded Up Robust Features. In: Leonardis A, Bischof H, Pinz A, editors. *Computer Vision – ECCV 2006* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006 [cited 2019 Oct 31]. p. 404–17. Available from: http://link.springer.com/10.1007/11744023_32
- [19] MORPH [Internet]. [cited 2019 Oct 31]. Available from: https://www.faceaginggroup.com/?page_id=529
- [20] Flanagan PA. Face Recognition Grand Challenge (FRGC) [Internet]. NIST. 2010 [cited 2019 Oct 31]. Available from: <https://www.nist.gov/programs-projects/facerecognition-grand-challenge-frgc>
- [21] Liu H, Lu J, Feng J, Zhou J. Group-aware deep feature learning for facial age estimation. *Pattern Recognition*. 2017 Jun 1;66:82–94.
- [22] FG-NET data by Yanwei Fu [Internet]. [cited 2019 Oct 31]. Available from: http://yanweifu.github.io/FG_NET_data/index.html

- [23] Hu Z, Wen Y, Wang J, Wang M, Hong R, Yan S. Facial Age Estimation With Age Difference. *IEEE Trans on Image Process*. 2017 Jul;26(7):3087–97.
- [24] Xin Geng, Chao Yin, Zhi-Hua Zhou. Facial Age Estimation by Learning from Label Distributions. *IEEE Trans Pattern Anal Mach Intell*. 2013 Oct;35(10):2401–12.
- [25] Geng X. Label Distribution Learning. arXiv:14086027 [cs] [Internet]. 2016 Apr 5 [cited 2019 Oct 31]; Available from: <http://arxiv.org/abs/1408.6027>
- [26] Chang K-Y, Chen C-S, Hung Y-P. A Ranking Approach for Human Ages Estimation Based on Face Images. In: 2010 20th International Conference on Pattern Recognition [Internet]. Istanbul, Turkey: IEEE; 2010 [cited 2019 Oct 31]. p. 3396–9. Available from: <http://ieeexplore.ieee.org/document/5597533/>
- [27] Lu J, Liong VE, Zhou J. Cost-Sensitive Local Binary Feature Learning for Facial Age Estimation. *IEEE Trans on Image Process*. 2015 Dec;24(12):5356–68.
- [28] Huerta I, Fernández C, Prati A. Facial Age Estimation Through the Fusion of Texture and Local Appearance Descriptors. In: Agapito L, Bronstein MM, Rother C, editors. *Computer Vision - ECCV 2014 Workshops* [Internet]. Cham: Springer International Publishing; 2015 [cited 2019 Oct 31]. p. 667–81. Available from: http://link.springer.com/10.1007/978-3-319-16181-5_51
- [29] Wang X, Guo R, Kambhmettu C. Deeply-Learned Feature for Age Estimation. 2015 IEEE Winter Conference on Applications of Computer Vision. 2015;534–41.
- [30] Xiaolong Wang, Kambhmettu C. Age estimation via unsupervised neural networks. In: 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG) [Internet]. Ljubljana: IEEE; 2015 [cited 2019 Oct 31]. p. 1–6. Available from: <http://ieeexplore.ieee.org/document/7163119/>
- [31] Criminisi A, Zisserman A, Van Gool LJ, Bramble SK, Compton D. New approach to obtain height measurements from video. In: Higgins K, editor. Boston, MA; 1999 [cited 2019 Oct 31]. p. 227–38. Available from: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=974201>
- [32] BenAbdelkader C, Yacoob Y. Statistical body height estimation from a single image. In: 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition [Internet]. Amsterdam, Netherlands: IEEE; 2008 [cited 2019 Oct 31]. p. 1–7. Available from: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4813453>
- [33] Guan Y-P. Unsupervised human height estimation from a single image. *JBiSE*. 2009;02(06):425–30.