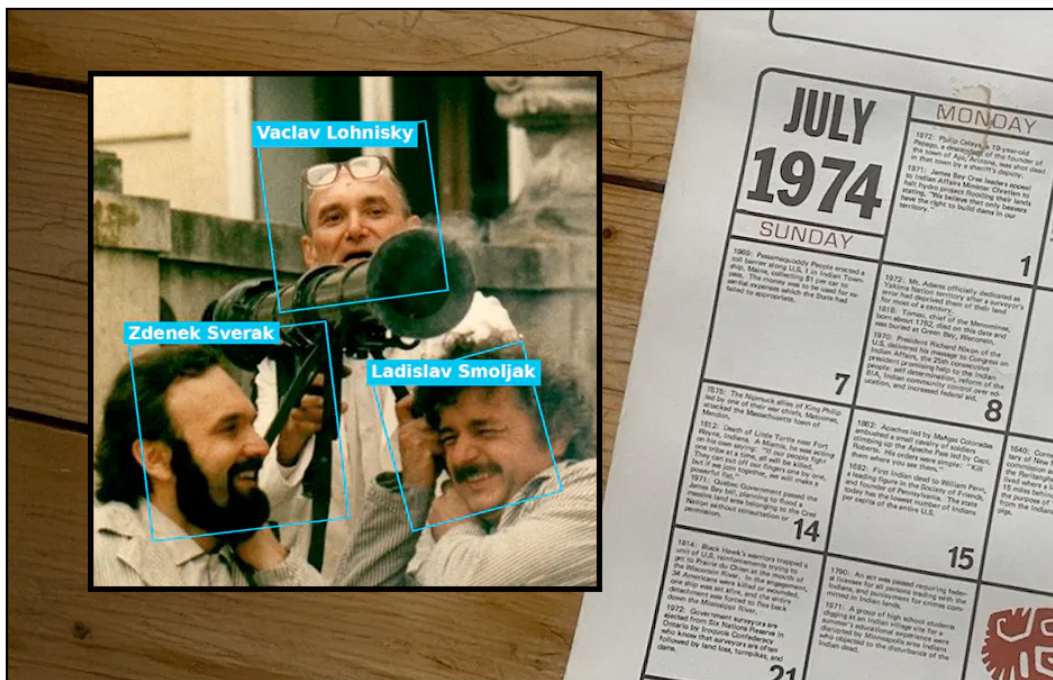


AI Can Guess a Photo's Year From People's Ages

Originally published at <https://www.unite.ai/ai-can-guess-a-photos-year-from-peoples-ages/>



New research shows AI can use people's faces to estimate the year a photo was taken, combining age guesses with known birth years to beat current scene-based methods.

Guessing the date of a photo used to be a fair bit easier than it is now, because hair and clothing fashions once evolved [at breakneck speed](#). For much-debated reasons, this churn of visual style [ended about thirty years ago](#), meaning that it is no longer quite so easy to look at a hairstyle or items of clothing and guess the year from this kind of visual clue.

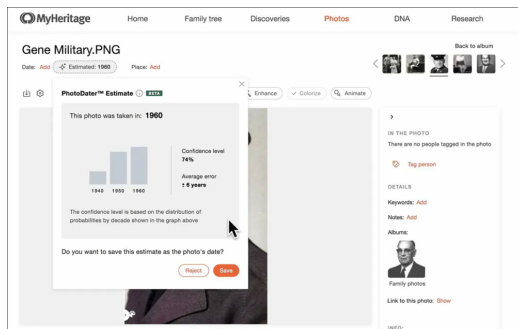
For some time, it was possible also to date images and movies based on the color resolution and [grain characteristics](#) of film stock. One did not have to be a forensic specialist; if you watched enough old movies, the cultural clues (such as music, cars, fashion, topics, etc.) would eventually become associated, by the viewer, with [film stock styles](#):



An illustration of the way that improvements in film stock gradually expanded the range of skin tones and lighting styles over time, moving from flat, frontal setups to more naturalistic and varied looks. [Source](#) (my own article)

An additional 'anchor' for dating a photograph was whether it was in black-and-white – an economy that became redundant after the popularization of digital photography early this century

A number of commercial and experimental systems, such as the MyHeritage subscription-bundled [PhotoDater](#) attempt to date photos using these and diverse other criteria.



An example of photo age estimation, from the MyHeritage PhotoDater subscription-only service. [Source](#)

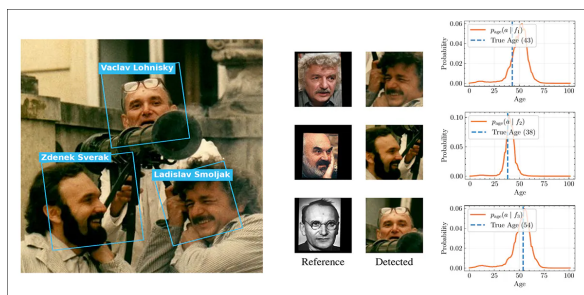
Absent other tell-tale signs, such as smartphones or other era-specific technology, the best way of telling the age of a photo taken in the last 15-25 years is if you are familiar with the person (i.e., a celebrity, or perhaps an acquaintance), and can estimate their *age*, which yields a rough equivalent year.

Facial Age as a Reference

In the field of computer vision, and in diverse other fields (i.e., forensics, archival processing, journalism, dataset architecture, etc.) the ability to determine the age of a photo is a prized goal, since many of the most interesting digital and analogue collections lack proper annotation and metadata, or even have incorrect metadata from previous (wrong) guesses.

Therefore it would be useful if an AI system could review photos the same way we do when looking back over our historical collections, and commenting ‘Oh yes, *that was when...*’. The question is, what could be the hook, absent the usual requisite clues?

A new research paper from the Czech Republic is offering an initial foothold into this approach, by exploiting AI-based *age recognition* systems, in concert with *facial* recognition systems linked to a common database of identities (in this case, an IMDB-style collection featuring Czech performers and film-makers):



A still from ‘Joachim, Put It in the Machine’ (1974), used to illustrate the dating process. The model detects known individuals in the photo, estimates their age using a facial age estimator (right column), and subtracts that value from each person’s birth year to generate a probability distribution over possible photo dates. The graphs show the likelihood of each age estimate, with dashed lines marking the person’s true age at the time of the photo. [Source](#)

The system works by detecting known individuals in a photo, estimating their facial age using a pretrained model, and subtracting this estimate from their documented birth year to generate likely dates for the photo. When multiple faces are present, the date estimates are aggregated to produce a final prediction.

The method was tested on images curated from the Czecho-Slovak Movie Database ([CSFD](#)), with the resulting approach, the authors assert, offering consistently better accuracy than scene-based models (static models that rely on background elements or visual context rather than faces) trained on the same data.

The schema for this method requires a central database that contains knowledge of a broad group of individuals, in this case the IMDB-style Czech movie database; but any similar collection the features confirmed birth dates and central date-confirmed events could yield a similar result.

The paper states:

‘Uniquely, our dataset provides annotations for multiple individuals within a single image, enabling the study of multi-face information aggregation. We propose a probabilistic framework that formally combines visual evidence from modern face recognition and age estimation models, and career-based temporal priors to infer the photo capture year.

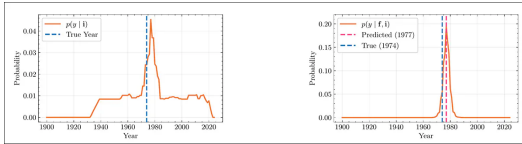
‘Our experiments demonstrate that aggregating evidence from multiple faces consistently improves the performance and the approach significantly outperforms strong, scene-based baselines, particularly for images containing several identifiable individuals.’

The [new paper](#) is titled *Photo Dating by Facial Age Aggregation*, and comes from two researchers at Czech Technical University in Prague, with the promise of a later code/data release.

Method

To estimate when a photo was taken, the authors’ new system looks at each detected face and attempts to guess who it might be, using the aforementioned database of known people. Since a person can only appear once in a photo, the system checks all combinations of possible identities and uses their known birth years to guess how old each person looks.

After this, it works backward to estimate the most likely year that would make those ages line up:

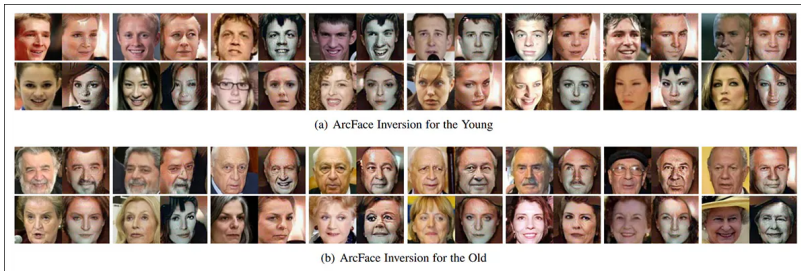


Left: the system builds a timeline showing when the recognized individuals were most active, based on their known careers.
Right: this is combined with facial age estimates to produce a final guess for when the image was taken.

To manage the great many possible identity combinations, the system assumes that faces are independent, and that each one's appearance depends solely on its identity and the date of the photo.

To estimate when a photo was taken, the system first guesses the age of each detected face using the NIST [cvut-002 model](#), which is based on a [ViT-B/16 architecture](#), and trained on a private dataset (which, the authors state, ranks highly in NIST's Face Analysis Technology Evaluation (FATE) [database](#)).

Once the person's birth year is known, the model converts the age estimate into a likely *photo* year by simply adding the age to the birth year, yielding a probability distribution over possible capture years. To assess how well a detected face matches a known identity, the system compares their [embeddings](#) in [ArcFace space](#):

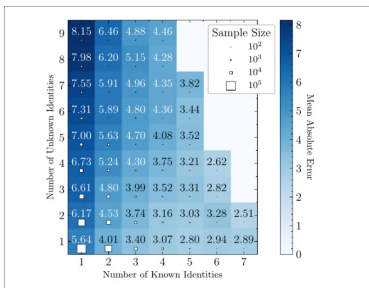


ArcFace, the central contributing architecture for the now-popular InsightFace model, was launched in 2015, destined to become an influential project in facial ass

Each identity is represented by an average embedding built from its reference portraits. The similarity between a test face and an identity is then measured using a [Von Mises Fisher Distribution](#), which models how tightly the identity's portraits cluster around that average. A shared sharpness parameter controls how confident the system is in those clusters, and is estimated using a [leave-one-out strategy](#) on the identity portraits.

The model defines five types of [priors](#) to estimate when a recognized person might appear in a photo: *uniform*; *decade*; *movie*; *image*; and a *convex combination prior* that mixes the strongest and weakest options, to test sensitivity to prior strength (i.e., the resilience of the priors under stress).

To handle faces that can't be confidently identified, the model includes a fallback 'unknown' identity with uninformative distributions, featuring a face likelihood that's flat in the embedding space, and a temporal prior flat across all years. This allows uncertain faces to be ignored without biasing the final date estimate:



How performance is affected when some faces in an image cannot be identified. Each square shows the average dating error for different numbers of known and unknown identities, with square size reflecting how common that combination is in the dataset. Error increases with more unknowns, but drops steadily as more known identities are added.

Data and Tests

The authors used the aforementioned CSFD dataset to furnish data for a new collection that they titled *CSFD-1.6M*. The dataset was built from scenes featuring several people, with each face labeled by identity and year. This structure was necessary to teach the model how faces relate to each other in context; single-face datasets such as IMDB-WIKI do not support this, since they label only one person per image.

Movie release years from the Czecho-Slovak Movie Database were used to estimate when each photo was taken, with each person in the image matched to a public profile featuring their birth year, and a portrait.

Subsequently, each face in the image was matched to one of the known identities, initially using ArcFace to create face embeddings, and computing an average embedding for each identity.

After this the [Hungarian algorithm](#) was used to assign faces to identities by comparing embedding similarity, with adjustments made when the number of faces detected via the [SCRFD-10GE](#) framework did not match the number of known individuals.

Statistic	Value
Total Scraped Images	1,493,357
Total Detected Faces	3,655,604
Faces Matched to an Identity	1,921,164
Final Faces (Matched & Dated)	1,653,871
Final Images	1,143,213
Total Annotatable Identities	75,952
Unique Identities in Final Set ⁴	46,523

Statistics from the CSFD-1.6M dataset, detailing scraped images, detected faces, identity matches, final annotated samples, and the available identity pool.

Matches were rejected if similarity was too low or if the estimated age differed too greatly from the known age, with greater tolerance allowed for older subjects, and faces were not filtered by quality or size.

The authors note the superiority of their curated set over that of the nearest comparable dataset, IMDB-WIKI:

‘Our dataset is not only substantially larger but, critically, consists of multi-person scenes required by our model. While no web-scraped dataset is free of label noise, our annotation pipeline leverages the explicit links between images and identity profiles provided by the database, aiming for higher-quality identity assignments.’

Their evaluation compared several versions of the dating system, to understand where its gains were coming from. One model assumed perfect knowledge of who was in the image, providing an upper bound on performance by removing any uncertainty in identity recognition, with the full version of the model then estimating identities and dates jointly, weighing different possible identity assignments before arriving at a final year estimate.

A simpler variant selected the single most likely identity configuration without marginalizing over alternatives, which proved nearly as effective in practice.

By contrast, the most basic baseline assigned each face independently and combined the resulting age-based year estimates, without considering whether the identities collectively made sense.

To test how much the method benefited from using faces at all, a separate model was trained to estimate the date directly *from the entire scene*. This scene-based model constitutes the strongest alternative approach currently used in image date estimation, since it can learn *era-specific* visual patterns across the full image, rather than relying on identity or age.

Metrics and Data

Mean Absolute Error (MAE) between the predicted year and the known ground truth was the central metric for the experiments.

The data was [divided](#) into five parts, with care taken to ensure that all images from the same movie were kept within a single partition. Three of these parts were used for training, one for validation, and one for testing. This five-fold rotation was applied to prevent [overfitting](#).

Since the face-based models were not trained on this dataset, no splitting was required, and instead, they were evaluated directly on the full CSFD-1.6M set.

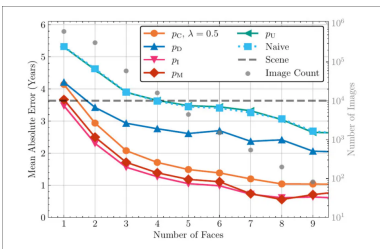
The *Scene* model was trained for 200 [epochs](#) under the [Adam](#) optimizer, with images resized to a 384×384 [crop](#).

Results

The results section of the paper is divided unusually across a number of performance indicators, with no single outstanding or central test. However, we’ll present a selection of the most pertinent results here.

The most important result is not a single number, but a pattern: facial aggregation models (especially the *Full* and *Top-1* variants) consistently outperform the strong *Scene* baseline whenever two or more known identities are present – even though the *Scene* model is trained directly on the dataset, supporting the central claim that identity-linked facial dating provides a more robust signal than holistic scene interpretation.

To evaluate the effect of temporal priors, the authors compared several configurations of their *Full* model. The strongest performance was obtained using the *Decade* Prior, which significantly outperformed both the *Naive* model (which uses no temporal prior) and the *Uniform* Prior (which assumes no preference over years):



Performance drops sharply for all methods as the number of faces increases, but models using realistic temporal priors such as the Decade Prior are affected far less. The Naive and Scene baselines remain flat or degrade with larger groups, while the Full model guided by informative priors maintains low error. The oracle-based priors, which rely on test-set statistics, define the lower bound on achievable performance.

To demonstrate the value of CSFD-1.6M beyond photo dating, the dataset was also tested as a pretraining resource for the broader task of facial age estimation. Following a standard evaluation protocol, [ResNet101](#) models were pretrained on CSFD-1.6M, and compared to counterparts pretrained on IMDB-WIKI and ImageNet. These models were then fine-tuned and evaluated across five popular benchmarks: [AgeDB](#); [AFAD](#); [MORPH](#); [UTKFace](#); and [CLAP2016](#).

Benchmark	Pre-training Dataset		
	ImageNet	IMDB-WIKI	CSFD-1.6M
AgeDB	7.05 $\pm$ 0.29	6.34 $\pm$ 0.25	5.25 $\pm$ 0.21
AFAD	3.19 $\pm$ 0.04	3.10 $\pm$ 0.03	3.04 $\pm$ 0.03
MORPH	2.98 $\pm$ 0.05	2.88 $\pm$ 0.07	2.76 $\pm$ 0.05
UTKFace	4.84 $\pm$ 0.08	4.64 $\pm$ 0.06	4.08 $\pm$ 0.03
CLAP2016	5.87	4.89	3.52

Mean absolute error (plus minus standard deviation) on five age estimation benchmarks, comparing models pretrained on ImageNet, IMDB-WIKI, and CSFD-1.6M. Lower values indicate better performance. CSFD-1.6M yields the strongest results across all benchmarks.

Across all five datasets, pretraining on CSFD-1.6M led to the lowest error rates, outperforming the other two pretraining sources by a clear margin – a performance gap that proved strongest on AFAD and CLAP2016, but remained consistent across the board.

We refer the reader to the rest of the somewhat fragmented results section in the source paper, which also deal extensively with ablation studies.

Conclusion

Though the new paper quickly becomes dense and unapproachable for the casual reader, the topic addressed is among the most interesting and relevant in computer vision literature – not least because it crosses over rather adroitly into anthropology and cultural studies, where the constants are hard to pin down.

* Just as musical evolution also [slowed its rate of change](#).

First published Monday, November 10, 2025

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More