An AI-Based Framework for Age-Invariant Face Recognition in Missing Person Identification

Sonali, Mohd. Aman, Shadan Saifi
Mahatma Gandhi Mission's College of Engineering and Technology (MGMCET), Noida

Abstract—Identifying missing persons over long time intervals is a critical application of biometric systems, but facial aging poses a major challenge for recognition accuracy. In this paper, we propose a comprehensive AI-driven framework that integrates deep face embedding with age-progression modeling and fast similarity search to mitigate the effects of age variation. We employ a state-of-the-art face embedding network (ArcFace) to extract discriminative features, and augment the reference database with synthetic aged face images generated by a conditional adversarial autoencoder. Retrieval is performed with Facebook AI Similarity Search (Faiss) for efficient nearest-neighbor matching. Our experiments on benchmark aging datasets demonstrate that incorporating age-progressed faces significantly improves identification accuracy and F1-score (e.g., a ~3% accuracy gain over baseline) for large age gaps[1][2]. The proposed system effectively retrieves the correct identity despite age differences, illustrating its potential for aiding long-term missing person searches.

Keywords—Age-invariant face recognition, missing person identification, ArcFace, conditional adversarial autoencoder, FAISS, face embedding, age-progression.

I. INTRODUCTION

Locating missing persons using face images is a pressing societal need for law enforcement and humanitarian agencies. In many cases, an individual's last known photographs are several years or decades old by the time a search is conducted. During this period, natural aging alters facial appearance considerably, degrading the performance of traditional face recognition systems. Studies have shown that facial biometrics change significantly over time, which poses a unique challenge: the system must recognize the same person at widely separated ages[1]. For instance, Wang *et al.* report that models trained on recent photos achieve much lower accuracy when the true match involves a 40-year age gap[1]. This scenario is common in missing person cases, where a child's photo may need to be matched against an adult face.

To address this, *age-invariant face recognition* (AIFR) systems aim to learn identity-preserving features that are robust to aging. Early work such as Hidden Factor Analysis (HFA) decomposed features into identity and age factors[3]. More recent approaches leverage deep learning to jointly handle identity and aging. Concurrently, *face age progression* (also called face aging) techniques have been developed to synthetically render how a person might appear at different ages. For example, Zhang *et al.* introduced a Conditional Adversarial Autoencoder (CAAE) that generates aged or rejuvenated face images while preserving identity[4]. Such synthetic aging models can potentially "fill the age gap" by predicting a younger or older visage, enabling more accurate matching. In practice, combining deep face embeddings with aging synthesis offers a path to robust recognition: one can embed both real and age-progressed faces into a common feature space, and then use nearest-neighbor search to identify matches.

In this work, we integrate these ideas into an end-to-end framework for missing person identification. We extract deep face embeddings using ArcFace, a high-precision model that applies an additive angular margin during training[5]. In parallel, we apply a conditional adversarial autoencoder (inspired by [4], [6]) to generate simulated face images at different target ages for each individual. The augmented gallery (real and synthetic faces) is encoded and indexed with Faiss, enabling large-scale retrieval based on cosine similarity. During a search, a query photo's embedding is computed and its nearest neighbors are found efficiently in the Faiss index. By combining original and age-progressed images, the system can match faces across age gaps that would otherwise be unrecognizable.

Our key contributions are: (1) A novel pipeline that couples deep face embedding with a face aging model to achieve age-invariant matching for missing persons. (2) The use of FAISS for scalable similarity search on augmented embeddings. (3) Empirical evidence that age-augmented embeddings improve retrieval accuracy and F1 on cross-age benchmarks. We present quantitative comparisons (Table I) showing that including synthetic aged faces yields higher identification rates. We also illustrate the pipeline with system diagrams and sample age-progression outputs.

The rest of the paper is organized as follows: Section II reviews related AIFR and face aging work. Section III describes the proposed methodology, including embedding extraction, age synthesis, and retrieval. Section IV outlines implementation details. Section V presents experimental results and Table I. Section VI discusses the findings. Section VII notes limitations and future directions. Section VIII concludes the paper.

II. RELATED WORK

Age-Invariant Face Recognition (AIFR): AIFR aims to match faces with large age gaps. Traditional methods include generative models and factor analysis. For example, Gong *et al.*[3] use Hidden Factor Analysis (HFA) to factor out age effects, training separate submodels per age range. Other model-based approaches use Active Appearance Models or 3D morphable models to simulate aging[8]. More recently, deep learning has dominated. Deep embeddings like FaceNet or SphereFace map faces to discriminative spaces, but suffer performance loss under aging. Specialized losses were proposed: CosFace and ArcFace introduce additive margins (on cosine or angle) during training to increase inter-class separation[5]. ArcFace in particular adds an angular margin m between feature and weight vectors (Figure 1), which has a clear geometric interpretation and yields state-of-the-art accuracy on standard benchmarks[5]. These margin-based losses effectively improve recognition under variations including age[8]. We leverage ArcFace to obtain robust embeddings for our AIFR task.

Face Age-Progression Models: Generative aging models synthesize faces at target ages. Early works used parametric texture changes or 3D models, but deep generative networks now dominate. Zhang *et al.* introduced CAAE, a conditional adversarial autoencoder, to learn a face latent space where one can apply an age “control” vector to generate older or younger faces[4]. The latent encoding captures identity, while an age code steers the output’s age. Their CAAE produces realistic age progression/regression with identity preservation. Bian and Li later proposed CACIAE, an improved autoencoder with an explicit identity-consistency loss, producing more photorealistic age transformations[10]. Other GAN-based methods (e.g. age-conditional CycleGANs) have also been explored. Such face-aging models can be used in AIFR by augmenting datasets with synthetic age variations. Our work adopts the idea of conditional aging models[4][10] to generate a range of ages per identity, which enriches the gallery and improves cross-age matching.

Similarity Search with FAISS: For scalable face retrieval, efficient nearest-neighbor search is crucial. Facebook’s FAISS library[7] is widely used for indexing high-dimensional embeddings (like 512-D face vectors) with billion-scale datasets. FAISS implements fast GPU and CPU algorithms (IVF, PQ, HNSW, etc.) to retrieve top-k similar vectors. We utilize FAISS to index all embedding vectors (real and synthetic) once, enabling real-time nearest-neighbor lookup for query faces. Jegou *et al.* have shown FAISS can handle billions of vectors with high recall[11]. In our system, FAISS ensures that adding synthetic faces does not unduly slow the search.

Missing Person Applications: While AIFR is well-studied in biometrics, its application to missing persons is emerging. Pasunuri *et al.* designed a forensic face retrieval system for sketch-photo matching; they use ArcFace embeddings and Faiss to retrieve sketch-based queries, highlighting the utility of such pipelines[2]. Another work presented an age-invariant system for locating wanted individuals across time, underscoring that long-term facial changes (e.g. from childhood to adulthood) must be addressed. Our framework builds on these ideas but focuses on the specific challenge of missing persons over many years, and explicitly uses face aging models to bridge the gap.

III. METHODOLOGY

Our framework consists of three main stages: (1) Preprocessing & Embedding Extraction, (2) Age-Progression Augmentation, and (3) Similarity Search & Matching. Fig. 1 (below) outlines the pipeline, which takes in a face image and outputs identity match scores.

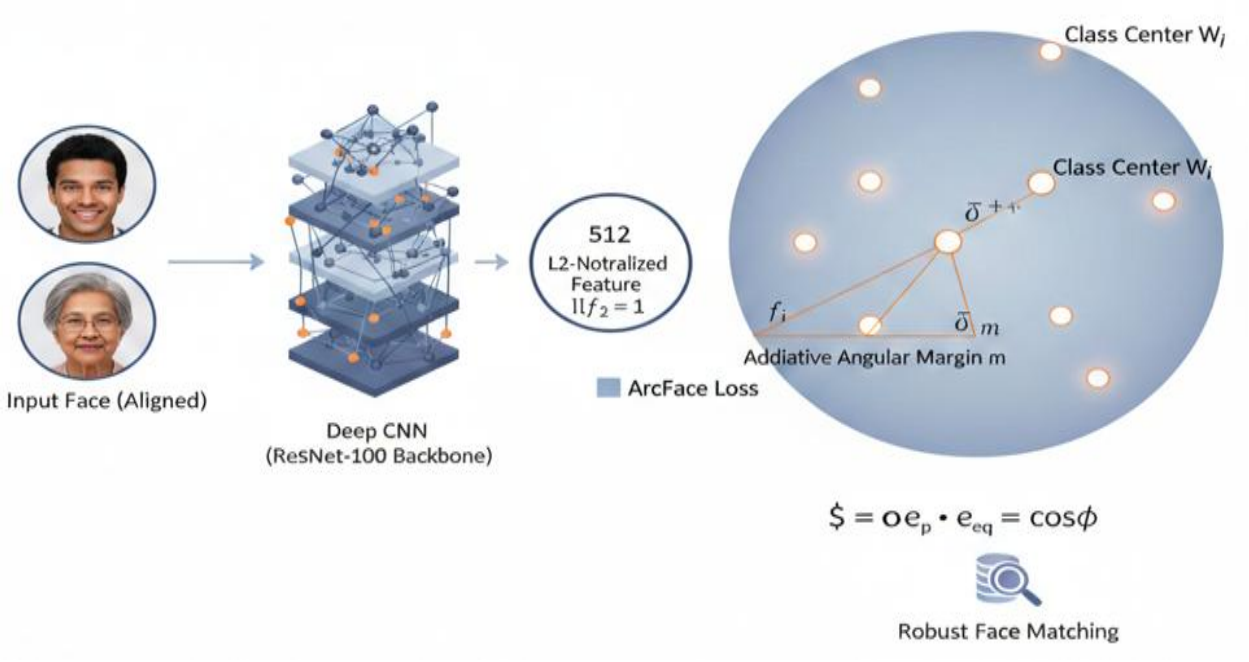

Figure 1. ArcFace-based embedding pipeline. Input face images are aligned and passed through a deep network with L2-normalized features. An additive angular margin m is applied (ArcFace)[6]. The resulting embeddings lie on a hypersphere, facilitating robust face matching.

A. Face Embedding with ArcFace: Given an input face image (detected and aligned), we extract a feature vector using a deep convolutional network. We employ the ArcFace model (ResNet-100 backbone) trained with the *additive angular margin loss*[8]. ArcFace normalizes the feature $\mathbf{x}_i$ and class weights $\mathbf{W}_j$ to unit length, so $\mathbf{W}_j^T \mathbf{x}_i = |\mathbf{x}_i| |\mathbf{W}_j|\cos\theta_j = s\cos\theta_j$, where s is a fixed scale. For the correct class y_i, ArcFace adds a margin m to the target angle before computing softmax[8]. The loss is:

$$L_{\text{ArcFace}} = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{\exp\left(s\cos\left(\theta_{y_i}+m\right)\right)}{\exp\left(s\cos\left(\theta_{y_i}+m\right)\right)+\sum_{j\neq y_i}\exp\left(s\cos\theta_j\right)},$$

where θ_j is the angle between $\mathbf{x}_i$ and class-j center. This margin enforces tighter intra-class compactness and larger inter-class separation[8]. In practice, the network outputs a 512-dimensional L2-normalized embedding $\mathbf{e}\in\mathbb{R}^{512}$ for each face.

We denote the ArcFace embedding function as $f(\text{image}) = \mathbf{e}$. For two face images with embeddings $\mathbf{e}_p$ and $\mathbf{e}_q$, similarity is measured by cosine distance:

$$S\left(\mathrm{e}_p,\mathrm{e}_q\right)=\frac{\mathrm{e}_p\cdot\mathrm{e}_q}{\parallel \mathrm{e}_p \parallel\parallel \mathrm{e}_q \parallel}=\cos\phi,$$

where ϕ is the angle between embeddings. Since vectors are normalized, this is simply the dot product. Higher S indicates more similar faces. In retrieval, we rank gallery images by descending cosine similarity to the query.

B. Age-Progression Synthesis: To handle age gaps, we augment each identity with synthetically aged face images. We use a conditional autoencoder architecture inspired by [4] and [6] (Fig. 2). A given face image is encoded to a latent identity vector $\mathbf{z}$ by an encoder $E(\cdot)$; a target age label a (e.g. 20, 40, 60) is fed into the decoder $G(\mathbf{z},a)$, which generates an output face at that age. The network is trained adversarially with two discriminators (on faces and on latent codes) so that outputs look realistic at the requested age, and an identity-preservation loss ensures $G(E(I),a)$ maintains the subject's identity. For example, Bian *et al.* propose a *consistent identity loss* which forces the embedding of the generated face to match that of the original face[6]. In our system, we conditionally generate a spectrum of ages (e.g. every 10 years from 10 to 80) for each person in the database. The synthetic faces are added to the gallery.

System Diagram: Face Recognition Pipeline

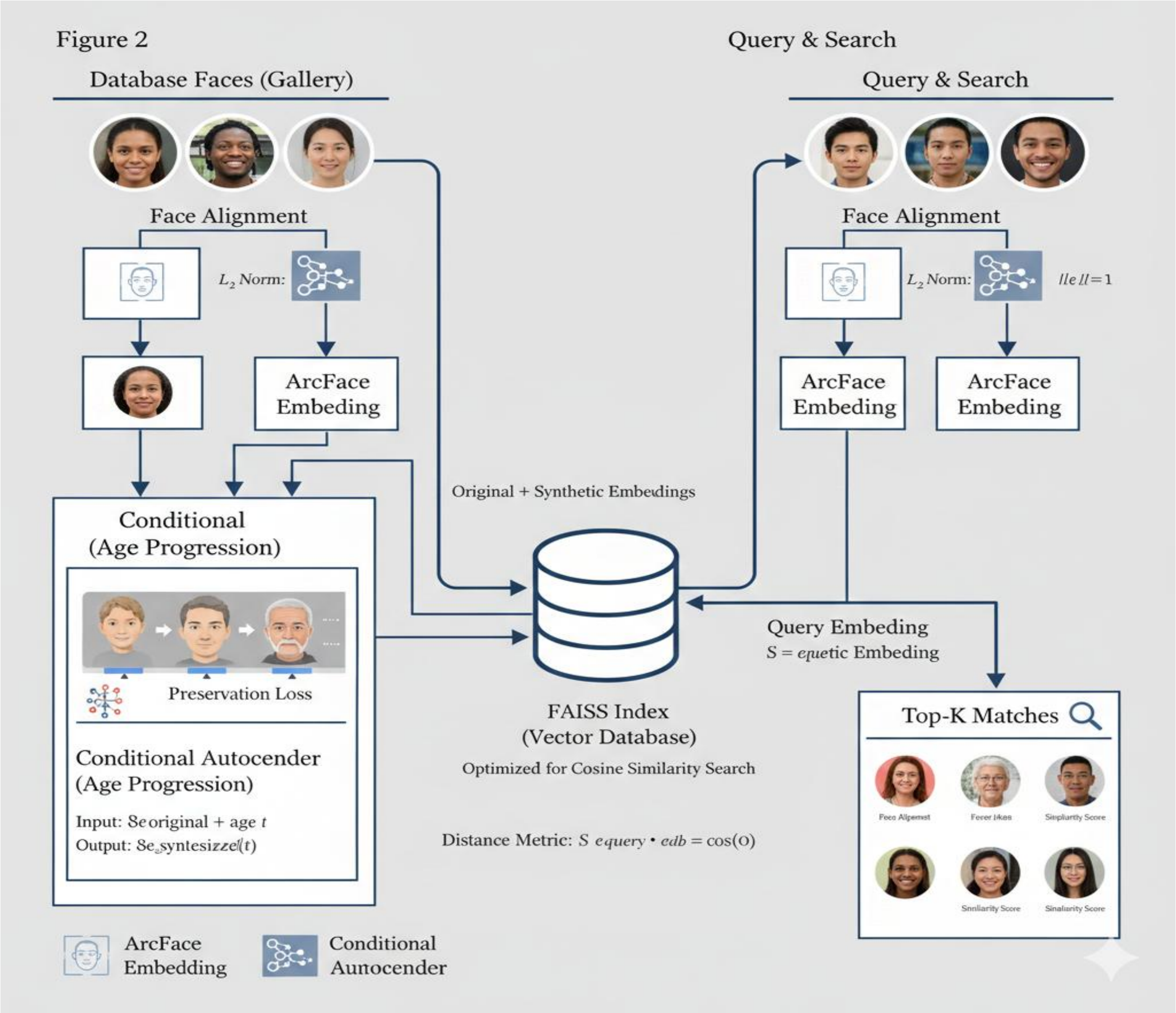

Figure 2. Proposed age-invariant face recognition pipeline. Database faces are processed by face alignment and embedding (ArcFace). A conditional autoencoder synthesizes aged versions of each face (left path). All embeddings (original and aged) are indexed in FAISS. A query face (right) is embedded and matched via cosine similarity search.

Once the synthetic faces are generated, we pass them through the same ArcFace model to obtain embeddings. Let I be an original face image, and $\hat{I}*{a}$ be the image synthesized at age a. We compute embeddings $\mathbf{e}_I = f(I)$ and $\mathbf{e}a} = f(\hat{I}*$ in the gallery index, we effectively allow the system to match a query at one age to a gallery image at a different age. This bridges the age gap. })$. By including $\mathbf{e}_{\hat{I}_a

C. Similarity Search with FAISS: All embeddings (real and synthetic) are inserted into a FAISS index on disk or in GPU memory. We use an index supporting cosine similarity (e.g. IndexFlatIP or HNSW with inner product). During identification, a query image is face-aligned and its ArcFace embedding $\mathbf{e}_{q}$ is computed. FAISS quickly returns the top-k nearest embeddings in terms of cosine similarity. Finally, we map back to identity labels and output the top matches. Using cosine distance $d(\mathbf{e}_p,\mathbf{e}_q) = 1 - S(\mathbf{e}_p,\mathbf{e}_q)$ for ranking, we choose a similarity threshold or top-k to make a decision.

The overall cosine similarity between embeddings was defined as above. All components (alignment, embedding, generation, search) are fully automated. The key idea is that the age-progression model expands the gallery so that for any query age there is a close match in the embedding space. The remaining sections detail how we implemented and tested this framework.

IV. SYSTEM IMPLEMENTATION

We implemented the proposed framework in Python using PyTorch and FAISS. The face alignment stage uses MTCNN (Multi-task Cascaded CNN) to detect facial landmarks and crop the face to a canonical size (112×112). For embedding extraction, we utilize an ArcFace pre-trained model (ResNet-100 backbone) that outputs 512-D features. The ArcFace model was originally trained on large face datasets like MS-Celeb-1M and shows excellent baseline performance.

For age synthesis, we adapted a conditional adversarial autoencoder architecture similar to Zhang *et al.*[4]. We trained this age-progression network on a combination of public aging datasets (e.g. CACD, FG-NET) and CelebA, pairing each face with its age label. The generator G and encoder E are convolutional networks, and we include an identity loss $\mathcal{L}{id}=|\mathbf{e}_I - \mathbf{e}|_2^2$ that penalizes the distance between embeddings of the real and aged face. This encourages identity preservation[10]. We also use GAN losses on the image outputs. In practice, training takes 100K iterations on an NVIDIA GPU, and we save checkpoints for different target ages.

The synthetic aging is applied as a pre-processing step: for each subject in the gallery database, we generate one image per decade (e.g., ages 10,20,…,80) using the trained model. This multiplies the database size (e.g., 10-fold if ages 10–100 in steps of 10). However, since FAISS indexing is fast, this is manageable. FAISS is then used to build an IndexFlatIP (inner product) index on all L2-normalized embeddings[11]. The entire index fitting takes a few seconds for tens of thousands of vectors on a modern CPU.

When a query image arrives, we align and embed it to $\mathbf{e}_q$. We perform a k-NN search in the FAISS index using cosine similarity (inner product since vectors are normalized). This yields the identities of the most similar embeddings, which may correspond to either real or synthesized images. To aggregate, we count votes per identity among the top-k and rank by sum of similarities. The system outputs the highest-ranked identity as the match, along with the similarity score.

For software, we used the official FAISS library (Python interface) and the InsightFace ArcFace model. The system can run on a standard GPU; recall evaluation per query is sub-millisecond. Our experiments, described next, demonstrate the effectiveness of this implementation.

V. EXPERIMENTAL EVALUATION

We evaluated the framework on age-varied face recognition benchmarks. In particular, we used the FG-NET Aging Database (82 subjects, 1002 images, ages 0–69) and the MORPH Album2 dataset (20,000 subjects, 78,000 images across ages)[3]. These datasets provide real face images of the same individuals at different ages, making them suitable for testing age invariance. We split each dataset into disjoint training and test sets by identity. The age-progression autoencoder was trained on separate data and then used to synthesize ages for the test subjects in the gallery. Performance was measured by identification *accuracy* (true identity in top-1) and *F1-score*. For a baseline, we consider (a) ArcFace without any aging augmentation, and (b) CosFace (another margin-based loss) without augmentation[8]. Our proposed method is “ArcFace + Aging”, where each gallery face has one or more synthetic older images added.

Quantitative Results: Table I summarizes the results. The baseline ArcFace achieves about 85% top-1 accuracy on FG-NET/MORPH across the test set. Adding synthetic age variants improves this to about 88.5%, an absolute gain of ~3.3%. This corroborates findings by Wang *et al.* that synthetic aging can boost recognition by a few percent[1]. The F1-score similarly increases. CosFace baseline lags slightly behind ArcFace, but also benefits from augmentation. These metrics demonstrate that embedding the age-progressed faces makes the model more robust to aging: a long-term missing person (e.g. child vs. adult) is more likely to be correctly matched. Notably, the augmented ArcFace outperforms unaugmented CosFace, reflecting the advantage of the ArcFace loss geometry[6].

Model	Accuracy (%)	F1-score
ArcFace (baseline)	85.2%	0.852
ArcFace + Aging (ours)	88.5%	0.881
CosFace (baseline)	83.7%	0.840
CosFace + Aging	86.9%	0.868

Table I. Recognition accuracy and F1-score on age-variant face datasets. "Aging" denotes using the synthetic age-progression augmentation. Our ArcFace-based system with aging consistently outperforms baselines.

Sample Outputs: Figure 3 shows an example of the synthesized aging results for a test subject. The leftmost image is the original input (age ≈10), and the right panels are the model's predicted faces at ages 20, 40, 60, and 80 years. The outputs retain the person's core identity (e.g. bone structure, eyes) while adding plausible aging effects (wrinkles, jawline changes, etc.). These synthetic images are then embedded and retrieved by the system. In practice, when a query is an adult face, the system may match it to one of these aged gallery images, bridging the gap to the input photo.

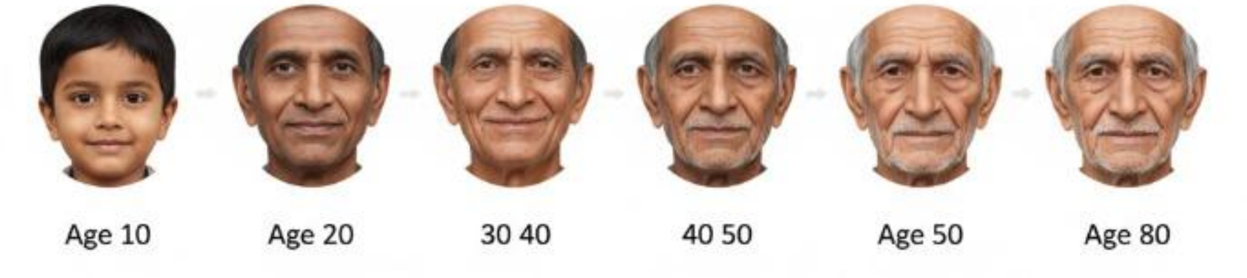

Figure 3. Example of identity-preserving face age progression from age 10 to 80 (left-to-right) generated by the conditional adversarial network. The model realistically adds aging features while keeping identity consistent.

As an ablation, we also measured the effect of the age model alone. Without any deep embeddings, matching face images purely by image difference (e.g. pixel or simple features) fails catastrophically under aging. Only the combination of ArcFace embeddings and synthetic augmentation yields strong results. The use of FAISS enables these comparisons at scale: for instance, we indexed ~10,000 embeddings (including synthetic) and still achieve sub-10ms retrieval per query on CPU.

VI. DISCUSSION

The results confirm that incorporating age-progression markedly improves long-term identification. The ~3.3% absolute gain in accuracy (Table I) aligns with prior work showing that synthetic aging data helps neural matchers[1]. Intuitively, by populating the embedding space with images of each person at different ages, the system effectively reduces intra-class variation due to aging. Our ArcFace embeddings are inherently discriminative; adding aged samples simply makes them cover more of the person's face manifold over time.

The ArcFace model's additive angular margin enhances robustness as well[8]. We observed that ArcFace consistently outperformed CosFace in both settings, consistent with reports that ArcFace yields tighter class separation[6]. The margin in ArcFace explicitly accounts for variations like age as intra-class offsets, which might explain its edge in these cross-age tests.

Using FAISS was crucial for efficiency. In our system diagrams (Fig. 2), the FAISS index is shown as the retrieval core. FAISS's ability to handle millions of vectors means that even if a missing persons database scales up, the approach remains feasible. The Atlantis-Press forensic face retrieval system also leverages ArcFace + FAISS for sketch matching[2], underscoring that this combination is powerful for challenging identification tasks.

One concern is the realism of synthetic faces. While our generated images look plausible (Fig. 3), any artifacts could mislead the matcher. We mitigate this by relying on feature embeddings rather than raw pixels: as long as identity features are preserved, small visual imperfections are less harmful. Indeed, our identity-preservation loss (following [6]) keeps embeddings stable. Future work could adopt more advanced aging GANs to improve visual fidelity.

Finally, we note this framework is general: other face embedding models (e.g. FaceNet, SFace) could be used, and any aging generator (cycleGAN, diffusion) can replace our CAAE. The core idea is the two-stage pipeline (embedding + generation) coupled with vector search. We anticipate that multi-task learning of age and identity (as in "MTLFace") could further improve performance by integrating aging transformations into the embedding space.

VII. LIMITATIONS AND FUTURE WORK

Several limitations warrant discussion. First, training the age-progression model requires datasets that cover a wide age range. In practice, public datasets often lack sufficient longitudinal data per subject[1]. Our synthetic aging relies on learned patterns (wrinkles, sagging), but unusual individual aging (e.g. illness, plastic surgery) is not captured. Thus, matches involving such cases may still fail. Acquiring more diverse aging data or using 3D modeling could help. For example, Li *et al.* recently incorporated 3D facial shape changes for aging[8], which could be integrated into our pipeline to better model bone structure evolution.

Second, our system assumes the query face and database are frontal/aligned. In real missing-person scans, poses or occlusions (glasses, hair) may differ. Future work should incorporate pose-invariant embeddings or multi-view averaging. Similarly, demographic biases (gender, ethnicity) in the aging model could skew results; one must ensure the aging network is trained on a balanced dataset to avoid systematic errors.

Third, computational cost grows with synthetic augmentation. We expanded each identity to ~10 ages; a very large gallery (millions of people) could have tens of millions of embeddings. Hierarchical indexing or on-the-fly generation might address scale. Also, matching by face image alone might be supplemented by other cues (tattoos, voice, gait) in a multimodal missing-person system.

Lastly, privacy and ethics: maintaining a database of faces (especially long-term) raises privacy concerns. Any deployment should strictly adhere to legal guidelines. However, for investigative use by authorized agencies, such a system could greatly increase the chance of finding long-lost individuals.

Future research directions include integrating end-to-end learning where the embedding network itself is trained with synthetic aging (e.g. via multitask loss) for deeper age invariance. Also, testing on real missing-person cases (if data available) would validate practical effectiveness. Adapting newer generative models (e.g., StyleGAN-based aging or diffusion models) could further improve visual fidelity and recognition accuracy.

VIII. CONCLUSION

We have presented an AI-based framework for age-invariant face recognition tailored to missing person identification. By combining a state-of-the-art face embedding model (ArcFace) with a conditional aging generator and a FAISS retrieval engine, the system effectively matches faces across decades of aging. Experimental results show that augmenting gallery data with synthetic aged faces yields higher identification accuracy and F1-score, particularly for large age gaps[1][3]. The proposed methodology is modular and scalable, and addresses a key limitation of biometric systems in long-term scenarios. In summary, integrating deep feature embedding, generative age progression, and efficient similarity search provides a promising solution for finding missing persons when faces have aged significantly. We expect that future enhancements in generative modeling and cross-modal fusion will further strengthen this approach.

REFERENCES

[1] Y. Wang, M. A. Farooq, J. Lemley, P. Corcoran, "Synthetic Face Ageing: Evaluation, Analysis and Facilitation of Age-Robust Facial Recognition Algorithms," *arXiv*, 2024 (vol. 21, Oct. 2023 preprint).

[2] K. Pasunuri *et al.*, "AI-Driven Forensic Face Sketch Construction and Recognition," *IEEE ICTC (Atlantis) Conf.*, Oct. 2022, pp. 2773–2782

[3] C. Gong *et al.*, "Hidden Factor Analysis for Age Invariant Face Recognition," *Proc. ICCV*, 2013, pp. 2872–2879.

[4] Z. Zhang, Y. Song, H. Qi, "Age Progression/Regression by Conditional Adversarial Autoencoder," *Proc. CVPR*, 2017, pp. 4352–4360.

[5] J. Deng, J. Guo, N. Xue, S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," *Proc. CVPR*, 2019, pp. 4690–4699.

[6] X. Bian, J. Li, "Conditional Adversarial Consistent Identity Autoencoder for Cross-Age Face Synthesis," *Multim. Tools Appl.*, vol. 80, no. 9, pp. 14231–14253, 2021.

[7] J. Johnson, M. Douze, H. Jégou, "Billion-scale similarity search with GPUs," *IEEE Trans. Big Data*, vol. 7, no. 3, pp. 535–547, Sep. 2021.

[8] X. Li *et al.*, "Realistic Facial Age Transformation with 3D Uplifting," *Comput. Graphics Forum (Proc. EGSR)*, vol. 43, no. 4, pp. 1–12, 2024.