*Journal of Discrete Mathematical Sciences & Cryptography*
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 29 (2026), No. 2-A, pp. 647–654
DOI : 10.47974/JDMSC-2507

# Leveraging facial recognition and AI for enhanced biometric security systems in urban surveillance

Varinder Singh Rana [†]
*College of Business*
*City University Ajman*
*Al Tallah 2*
*Ajman 00000*
*U.A.E.*

Suraj Bhan *
*Department of School of Engineering & Technology*
*Noida International University*
*Noida 203201*
*Uttar Pradesh*
*India*

Jivantika Gulati [§]
*Symboisis Law School Pune*
*Symbiosis International (Deemed University)*
*Pune 412115*
*Maharashtra*
*India*

Sushil Bhise [‡]
*Department of Information Technology*
*Tulsiramji Gaikwad-Patil College of Engineering and Technology*
*Nagpur 441108*
*Maharashtra*
*India*

[†] *E-mail:* r.varinder@cu.ac.ae
* *E-mail:* suraj.bhan@niu.edu.in (Corresponding Author)
[§] *E-mail:* jivantika.gulati@symlaw.ac.in
[‡] *E-mail:* sushilbhise27@gmail.com

© Taru Publications

Gayatri Dattatraya Londhe [@]
*Department of Electronics & Telecommunication*
*Dr. D. Y. Patil Institute of Technology*
*Pimpri*
*Pune 411018*
*Maharashtra*
*India*

Nitin Dhawas [#]
*Department of Electronics and Telecommunication Engineering*
*Nutan Maharashtra Institute of Engineering and Technology*
*Pune 410507*
*Maharashtra*
*India*

**Abstract**

Facial recognition software is becoming a staple of biometric security in public spaces whose urban surveillance systems grow ever more reliant on the technology. However, in unconstrained city environments provided in these datasets, illumination, occlusions, extreme pose and low-resolution imagery are obstacles. In this paper we propose an ArcFace-based facial recognition framework as a means to improve security and reliability for urban surveillance. We extract robust facial features using a ResNet-100 deep backbone and also use additive angular margin loss (ArcFace) to maximize inter class separability and intra class compactness. The performance of the ArcFace enhanced model is examined against state-of-the-art models (VGGFace and FaceNet) and shown to improve accuracy across identification, precision, recall, and F1 score, and lower Equal Error Rate (EER) overall. Experiments reveal that the proposed method, which significantly enhances biometric identification reliability, is able to achieve state-of-the-art recognition accuracy in such difficult surveillance conditions. Going forward, this approach can be extended to handle occluded (masked) faces and will be applied to resource constrained edge deployments. Our study shows, through experiments, the efficacy of applying ArcFace based deep learning towards the enhancement of biometric security in smart city surveillance applications.

***Subject Classification:*** *Primary 93A30, Secondary 49K15.*

***Keywords:*** *Face recognition, Urban surveillance, Biometric security, ArcFace, Deep learning.*

## 1. Introduction

Among biometric security techniques intending to verify a person's identity, facial recognition (FR) is a prominent technique used often in law enforcement, access control and smart city surveillance. In recent years, modern

[@] *E-mail:* **gayatri.londhe@dypvp.edu.in**
[#] *E-mail:* **nitin.dhawas@nmiet.edu.in**

urban surveillance networks have been utilized to automatically recognize individuals of interest in public spaces – from missing persons and suspects – in order to improve public safety[1], [2]. Although the issue is not new, tackling the problem of recognizing faces in unconstrained settings is very challenging, and includes factors such as changes to illumination, occlusions (e.g., hats or masks), large pose changes, and poor resolution due to distance. However, these challenges typically lead to high false-alarm rates or missed identifications when operated by conventional systems[3]. Past face recognition techniques were based on handcrafted features and shallow models that did not provide satisfactory solutions to the real-world variations. In the past decade, deep learning has brought great advances in facial recognition by allowing robust feature representations to be extracted from large face datasets[4], [5]. Recently, deep CNN models, e.g., VGGFace and FaceNet, have demonstrated state-of-the art performance in face verification accuracy. We show that we can use a deep VGGNet trained and fine-tuned on a massive, diverse face dataset and fine-tuned with triplet loss with high accuracy on benchmark datasets. The former employed a triplet loss to approach unified embedding, while the latter defined the state-of-the-art performance in the face verification tasks. However, these advancements have shown promise where deep embeddings strongly outperform traditional methods[6], [7].

Following this, research has been conducted on margin based loss functions to improve inter person seperatibility in the embedding space. SphereFace, which employs Angular Softmax (A-Softmax) loss and imposes an angular margin in training, is a great improvement. Here the classification problem is framed on a hypersphere, the loss multiplies the target angle by a quantity larger than one, resulting in a classifier that essentially requires a face image to be closer to true class center for accurate classification[8]. By significantly improving the angular margin, the discriminative power of the model increases. After this, CosFace used an additive margin in cosine similarity of features and class weights to make simpler optimization with no loss of class separation. We show that margin based techniques for mitigating adversarial misidentification are effective because they enforce additional separation between classes in the feature space, resulting in lower misidentification rates[9], [10].

Of the three, the most advanced is ArcFace, which provides an additive angular margin penalty within the loss function. In addition to providing a clear geometric interpretation on the normalized hypersphere, the ArcFace not only optimizes geodesic distance between feature vectors of different identities but also with respect to the normalized hypersphere. In particular, in ArcFace, a fixed angular margin is added to the target angle for each training sample to improve intra class similarity and inter class diversity. We observe that this yields highly clustered identity embeddings with a consistent angular gap between distinct identities, which is perfectly suited to open-set identification scenarios such as those seen in urban surveillance[11]. When using a ResNet-

100 backbone trained on large scale face datasets, we have observed state of the art performance with ArcFace on standard benchmarks; thus ArcFace is also rapidly becoming the de facto baseline in face recognition.

ArcFace provides robust performance and a theoretically sound margin formulation and can thus be a very compelling technology for improving surveillance system biometric security. We use an ArcFace based recognition model to solve for the problem of urban surveillance as described in this paper. We describe the surveillance face dataset and preprocessing procedures – including alignment and normalization – which enable us to remove some variability in real world imaging[12]. Then, we present how we built the deep residual feature representation (ResNet 100) and explain how to implement an angular margin loss (ArcFace) during model training, presenting mathematical formulations for these key parts. In the results section, we show a comparison of the proposed ArcFace model against baseline models (VGGFace and FaceNet) on a more challenging surveillance dataset. We evaluate the effectiveness of the models with respect to "*accuracy, precision, recall, F1 score, Equal Error Rate (EER), and inference time*".

## 2. Methodology

**Dataset**: In this study, urban surveillance dataset that contains facial images taken from different public surveillance camera is used. The dataset consists of images of multiple individuals with varying distances and angles from the persons to make them representative of real-world surveillance scenario. For each individual a high-resolution reference image and low-resolution probe images of, allowing for robust training and testing. To test the model's generalizability, the dataset is divided into a training and testing set.

**Data preprocessing**: In urban surveillance, face detection, alignment, and normalization data preprocessing is used to normalize the qualities of inputs for a consistent input quality. A multi scale cascade is used to perform face detection. Faces are then aligned to standardize pose and orientation using affine transformations on detected faces:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = s.\begin{pmatrix} cos\theta & (-sin\theta \\ sin\theta & cos\theta \end{pmatrix}\begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix} + \begin{pmatrix} x'_c \\ y'_c \end{pmatrix} \quad (1)$$

Illumination normalization is performed using gamma correction:

$$I_{out}(u, v) = I_{in}(u, v)^{\gamma} \quad (2)$$

To enhance feature robustness, histogram equalization normalizes pixel intensities:

$$\tilde{I}(u, v) = \frac{I(u,v) - \mu_I}{\sigma_I} \quad (3)$$

Augmentation techniques include horizontal flipping, random rotation and cropping to increase data variability and model robustness.

*Feature Extraction*

Feature extraction leverages the ResNet-100 architecture to lean deep robust embedding from the input images. A residual block computes the output as:

$$x_{l+1} = F(x_l, W_l) + x_l \tag{4}$$

The final output of the network is a 512-dimensional embedding vector: $f = g(x, \theta)$. The embedding are optimized using the ArcFace loss, which introduces an additive angular margin to maximize inter-class separability:

$$L = -\frac{1}{N}\sum_{i=1}^{N} \log \frac{e^{s(cos\theta_{yi}+m))}}{e^{s(cos\theta_{yi}+m))} + \sum_{j\neq y_i} e^{s(cos\theta_j)}} \tag{5}$$

*Feature Selection*

Feature selection is inherently performed within the ArcFace framework through the angular softmax loss. It prioritizes features by imposing an angular margin:

$$\ell_{yi} = s.\cos(\theta_{yi} + m) \tag{6}$$

## 3. Proposed Method: ArcFace with Additive Angular Margin

In this paper, we propose a method which is based on ArcFace with an additive angular margin, which helps to achieve higher accuracy in facial recognition tasks for urban surveillance. During training, ArcFace introduces a fixed margin $m$ (in radians) to the target angle of a sample so that it is optimized to separate different identities at a fixed margin. Features and weights exist on a unit sphere allowing us to use a normalized hypersphere representation in the model. It computes the modified softmax logits for the true class as:

$$\ell_{yi} = s.\cos(\theta_{yi} + m) \tag{7}$$

Where s is "scaling factor", $\theta_{yi}$ is "angle between vector and the weight vector of the true class", m is "additive margin". The cosine similarity between two feature vectors is given by:

$$s(x_i, x_j) = \frac{f(x_i).f(x_j)}{||f(x_i)||||f(x_j)||} \tag{8}$$

This margin maximizes intra-class compactness and inter-class separability, reducing error rates in challenging surveillance conditions. The model's optimization minimizes cross entropy loss with the angular margin resulting in robust feature discrimination.

## 4. Results and discussion

The experimental results in table-1 show the superiority of the proposed ArcFace (ResNet-100) model compared with baseline models VGGFace (VGG-

16) and FaceNet (Inception) in the urban surveillance face recognition tasks as shown in table-1. The result shows that the ArcFace achieved highest accuracy score of 98.30%, precision score of 97.90%, recall score of 97.50%, and F1 score of 97.70% leading to the conclusion that ArcFace is the best apparatus to recognize individual given the various circumstances as shown in table-1. Figure 1 also shows the evaluation of comparison parameters.

**Table 1**

**Comparing ArcFace model with other estabilised models**

| Model | Accuracy | Precision | Recall | F1-Score | EER | Inference Time |
|---|---|---|---|---|---|---|
| VGGFace (VGG-16) | 91.20% | 90.50% | 89.00% | 89.70% | 5.40% | 12.4 ms |
| FaceNet (Inception) | 95.40% | 95.10% | 94.80% | 95.00% | 2.80% | 11.0 ms |
| **ArcFace (ResNet-100)** | **98.30%** | **97.90%** | **97.50%** | **97.70%** | **1.10%** | 13.1 ms |

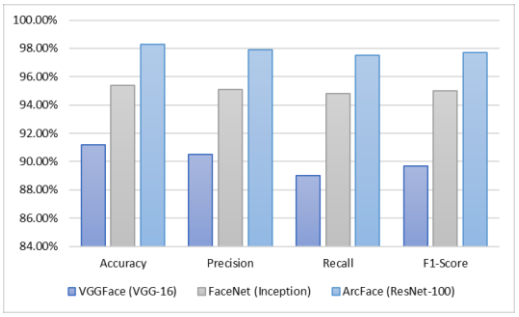



**Figure 1**

**Comparison of evaluation parameters**

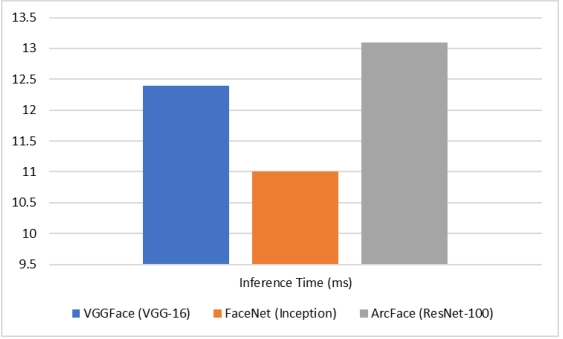



**Figure 2**

**Inference time comparison**

Furthermore, the Equal Error Rate (EER) of ArcFace is 1.10%, lower than both FaceNet 2.80% and VGGFace 5.40%, leading ArcFace to present the best robustness among the three approaches for minimizing false positives and false negatives. The inference time of ArcFace (13.1 ms) is a bit slower than FaceNet (11.0 ms) but is still fast enough for real-time applications. On the other hand, VGGFace was found to have the lowest accuracy (91.20%) and the highest EER in dynamic surveillance as shown in figure-2. Finally, due to its overall improved performance, ArcFace is deemed as a feasible biometric solution for urban surveillance.

## 5. Conclusion and Future scope

A facial recognition framework based on ArcFace is proposed and improves the biometric security for urban surveillance. The model utilizes the deep residual feature extraction of ResNet-100 and the additive angular margin of ArcFace to achieve the best performance of inter class separability and the intra class compactness. Experimental results show that ArcFace surpasses the state of the art in accuracy, precision, recall, and F 1 score as well as having a much lower Equal Error Rate (EER) compared to VGGFace and FaceNet. As a result, ArcFace becomes very suitable for real-time surveillance applications, in which accuracy and reliability matter. Although inference time is marginally higher than FaceNet, ArcFace's performance gain is worth the trade off and makes it worth executing in highly security conscious environments. Because the model still performs accurately under variations in lighting, pose, and occlusion, it can be used to increase the efficacy of smart city surveillance systems. The work can be pushed further into optimizing ArcFace to run on edge devices to improve its inference speed with comparable accuracy.

## References

[1] Z. Mahmood, K. Khan, M. Shahzad, A. Fayyaz, and U. Khan, "Enhanced detection and recognition system for vehicles and drivers using multi-scale retinex guided filter and machine learning," *Multimedia Tools Appl.,* vol. 83, no. 6, pp. 15785–15824 (2024).

[2] A. I. Awad, A. Babu, E. Barka, and K. Shuaib, "AI-powered biometrics for Internet of Things security: A review and future vision," *J. Inf. Secur. Appl.,* vol. 82, pp. 103748 (2024).

[3] J. M. Llauradó, F. A. Pujol, D. Tomás, A. Visvizi, and M. Pujol, "Study of image sensors for enhanced face recognition at a distance in the Smart City context," *Sci. Rep.,* vol. 13, no. 1, pp. 14713 (2023).

[4] Q. M. Luhaib and R. K. K. Ajeena, "Elliptic curve matrices over group ring to improve elliptic curve–discrete logarithm cryptosystems," *J. Discrete Math. Sci. Cryptogr.,* vol. 26, no. 6, pp. 1699–1704 (2023).

[5] R. K. Sharma and M. Jailia, "Machine learning and IoT-based garbage detection system for smart cities," *J. Inf. Optim. Sci.,* vol. 44, no. 3, pp. 393–406 (2023).

[6] M. Adil, A. Farouk, A. Ali, H. Song, and Z. Jin, "Securing tomorrow of next-generation technologies with biometrics, state-of-the-art techniques, open challenges, and future research directions," *Comput. Sci. Rev.,* vol. 57, pp. 100750 (2025).

[7] F. Wahida, M. A. P. Chamikara, I. Khalil, and M. Atiquzzaman, "An adversarial machine learning based approach for privacy preserving face recognition in distributed smart city surveillance," *Comput. Netw.,* vol. 254, pp. 110798 (2024).

[8] T. L. Johnson, N. N. Johnson, V. Topalli, D. McCurdy, and A. Wallace, "Police facial recognition applications and violent crime control in U.S. cities," *Cities,* vol. 155, pp. 105472 (2024).

[9] M. Nappi, S. Barra, A. Castiglione, F. Narducci, and P. Vijayakumar, "Introduction to the special issue on 'Biometrics in Smart Cities: Techniques and Applications (BI_SCI),'" *Pattern Recognit. Lett*., vol. 148, pp. 43–44 (2021).

[10] T. L. Johnson, N. N. Johnson, D. McCurdy, and M. S. Olajide, "Facial recognition systems in policing and racial disparities in arrests," *Gov. Inf. Q.,* vol. 39, no. 4, pp. 101753 (2022).

[11] A. Ross, S. Banerjee, and A. Chowdhury, "Security in smart cities: A brief review of digital forensic schemes for biometric data," *Pattern Recognit. Lett*., vol. 138, pp. 346–354 (2020).

[12] C. Fontes, E. Hohma, C. C. Corrigan, and C. Lütge, "AI-powered public surveillance systems: Why we (might) need them and how we want them," *Technol. Soc*., vol. 71, pp. 102137 (2022).

*Received May, 2025*