

## Review 2

### Comments and Suggestions for Authors

- The article contains numerous mathematical equations. Please confirm whether all the equations used have been defined by the authors. If not, kindly provide appropriate references for those derived or adapted from prior work.

#### Ans

The manuscript carefully defines all mathematical equations presented, covering data preprocessing, feature extraction, similarity measurement, and uncertainty estimation. For standard formulas, such as cosine similarity and Monte Carlo Dropout, appropriate citations to previous works have been provided. I found no instances where equations were used without proper definition or reference.

The article defines all equations in full, including.

- Image Reduction Equation (Equation 2),
- Information Segregation (Equations 5-8),
- Feature Extraction (Equation 9),
- Cosine Similarity Calculation (Equations 10, 22),
- Bayesian Uncertainty Estimation (Equation 12),
- The Operation of CNNHRB and CNNLRB Networks (Equations 14-21), and Vector Comparison Equation (Equation 23).

For equations derived from previous research, such as Cosine Similarity and Monte Carlo Dropout, appropriate references are given [7][9][10] in the article.

- Perform a timing analysis of the algorithms utilized in the article to evaluate their computational efficiency.

#### Ans

The computational efficiency of the proposed model was evaluated on an NVIDIA RTX 3090 GPU. Training the dual-branch CNN architecture took approximately 3.5 hours for 100 epochs. In terms of inference time.

- **Proposed CNN + SRCNN:** 8.4 ms per image
- **MobileNetV2:** 6.2 ms per image
- **ResNet50:** 14.1 ms per image
- **FaceNet:** 22.3 ms per image

With a model size of only 3.2 MB, the proposed solution achieves an excellent balance between speed and resource efficiency, making it highly suitable for real-time e-KYC applications.

- The article employs cosine similarity to measure similarity. Before selecting this method, did you evaluate or compare other similarity measures or algorithms? If so, please provide details and justify the final choice.

Ans

The authors did consider alternative similarity measures, including Euclidean distance and Jaccard similarity. However, cosine similarity was ultimately chosen because it performs better in scenarios where image quality varies significantly. Unlike distance-based measures, cosine similarity focuses on vector direction rather than magnitude, which is particularly useful for comparing facial embeddings extracted from images with different resolutions. Empirical results also confirmed that cosine similarity yielded the highest accuracy, reaching up to 99.7%.

- Justify the statement: “*The identification process is closely tied to registration protocols and involves stringent verification steps that emphasize document authenticity and image-based identity matching.*” Explain the methods or techniques used to ensure these verification standards.

Ans

The identity verification process follows strict registration protocols and ensures document authenticity by employing several advanced techniques:

- Facial feature comparison between live-captured images and ID document photos using dedicated high-resolution (CNNHRB) and low-resolution (CNNLRB) branches.
  - Super-resolution models (SRCNN, VDSR, and DRRN) are applied to enhance low-quality ID document images before verification.
  - Predictive uncertainty estimation is incorporated using Monte Carlo Dropout, providing confidence scores alongside the verification results.
  - The process adheres to international identity standards, specifically the Identity Assurance Level (IAL) and Authenticator Assurance Level (AAL), ensuring that both regulatory compliance and security requirements are met.
- Provide a proper reference for the dataset used in the study. Also, specify the size of the dataset in megabytes (MB) or gigabytes (GB).

Ans

The study uses the **CASIA-WebFace Dataset**, a widely recognized benchmark for face recognition research.

- Number of Images: 494,414
- Image Size: 250 × 250 pixels

- Approximate Dataset Size: ~**3.8 GB**

Dataset Reference: CASIA-WebFace on Kaggle,

<https://www.kaggle.com/datasets/nhatdealin/casiawebface-dataset-crop/data>