ASEC 2024 Conference

The 5th International Electronic Conference on Applied Sciences

04–06 December 2024 | Online

MDPI

A comparative analysis of metric combinations in face verification with machine learning techniques

Universitat d'Alacant
Universidad de Alicante

valgrAI Valencian Graduate School and Research Network of Artificial Intelligence

GENERALITAT VALENCIANA
Conselleria de Educación, Cultura, Universidades y Empleo

Héctor Penadés, Felix Escalona, Miguel Cazorla

Institute for Computer Research, University of Alicante, Alicante, Spain

INTRODUCTION & AIM

Face verification is essential in computer vision applications, including security, surveillance, and biometrics. Common metrics, such as L1, L2, and Cosine Similarity, are frequently employed to compare facial embeddings. However, each metric has limitations when it comes to generalizing across complex conditions, such as varying lighting, pose, or facial expressions. Occlusions and low image quality further complicate this task, underscoring the need for more robust solutions.

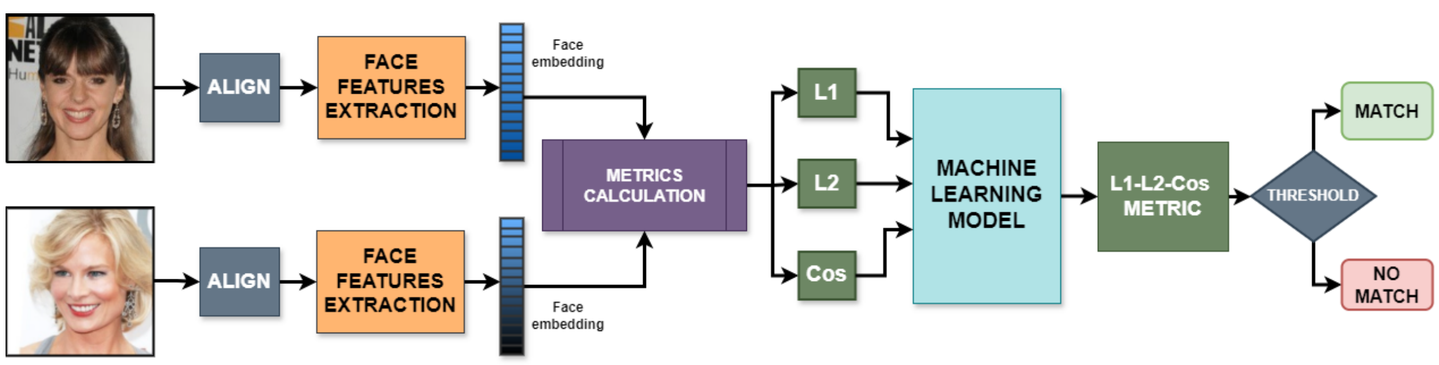
Objective

The study focused on improving face verification accuracy by combining embeddings obtained through machine learning models.

METHOD

Datasets

- Training dataset: *CASIA-WebFace*, a widely used dataset for face recognition model training.
- Evaluation dataset: *BUPT-BalancedFace*, chosen to ensure demographic balance and generalizability.



Feature extraction

- CNN-based models: AdaFace [1] and ArcFace [2] models were chosen for their high performance in face verification tasks.
- Transformer-based models: Swin transformers were employed to explore the effectiveness of transformer architectures.

Embedding metrics:

- Individual metrics: L1, L2, and Cosine Similarity were computed for facial embedding comparison.
- Metric combinations: Machine learning models were used to combine metrics, including Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), LightGBM, and XGBoost.

Model training:

Machine learning models were trained to learn an optimal combination of metrics to maximize accuracy in identifying whether two faces match.

Acknowledgments: Developed under the framework of the CIPROM/2021/17 Prometeo project entitled "Meebai: Una metodología para la educación consciente de las emociones basada en la inteligencia artificial", the Grant CIGE/2022/170 funded by Conselleria de Educación, Universidades y Empleo of Generalitat Valenciana, and by ValgrAI – Valencian Graduate School and Research Network for Artificial Intelligence and Generalitat Valenciana.

RESULTS & DISCUSSION

Metric	Accuracy (Adaface Model)
L1 Distance (L1)	92.5%
L2 Distance (L2)	92.4%
Cosine Similarity (Cos)	92.6%
Logistic Regression (LR)	93.1%
K-Nearest Neighbors (KNN)	94.0%
Support Vector Machine (SVM)	93.1%
Extreme Gradient Boosting (XGB)	93.4%
Gradient Boosting (GB)	93.4%

AdaFace

- ❑ This model achieved the best results, outperforming the transformer-based SwinFace model, which did not exceed 83% accuracy in any test.
- ❑ Cosine Similarity (alone) achieved 92.6% accuracy.
- ❑ L1-L2-Cosine Similarity combined with KNN improved accuracy to 94.0%, representing a 1.4 percentage point increase.
- ❑ These results highlighted the advantage of combining metrics for face verification.

ArcFace

- ❑ Accuracy increased from 86.5% to 87.3% when combined with LR. However, it remained less effective than AdaFace.

CONCLUSION

This study showed that combining multiple embedding metrics with machine learning models can improve face verification accuracy. While Cosine Similarity is a strong metric on its own, combining it with other metrics using models like KNN, LightGBM and XGBoost provided a significant improvement in accuracy. In addition, our findings indicate that CNN-based models outperform vision transformers for this specific task.

FUTURE WORK

Explore the combination of embeddings generated by different models, to leverage the strengths of each architecture and achieve higher accuracy in face verification.

REFERENCES

[1] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," *Proc. of CVPR*, 2019.

[2] S. Serengil and A. Ozpinar, "A Benchmark of Facial Recognition Pipelines and Co-Usability Performances of Modules," *Journal of Information Technologies*, vol. 17, no. 2, pp. 95-107, 2024. doi:10.17671/gazibtd.1399077.