Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 46 (2025), No. 1, pp. 43-52
DOI : 10.47974/JIOS-1850

Leveraging convolutional neural networks for real-time student attendance tracking

Rajesh Yadav *
Swati Gupta †
Department of Computer Science and Engineering
K. R. Mangalam University
Gurugram
Haryana
India

Meenakshi Malik §
Department of Computer Science and Engineering
BML Munjal University
Gurugram
Haryana
India

Poonam Yadav ‡
Department of Computer Science and Engineering
K. R. Mangalam University
Gurugram
Haryana
India

Abstract

This study presents an advanced Convolutional Neural Network (CNN) for facial detection and identification, outperforming traditional models. The method uses a channel-

* *ORCID-ID:* 0009-0003-2688-7631
† *ORCID-ID:* 0000-0002-9450-6381
§ *ORCID-ID:* 0000-0001-9118-8819
‡ *ORCID-ID:* 0009-0004-4386-4172

* *E-mail:* Himank9d2006@gmail.com (Corresponding Author)
† *E-mail:* swattiguptta@gmail.com
§ *E-mail:* meenakshimalik16@gmail.com
‡ *E-mail:* poonamy0312@gmail.com

© Taru Publications

wise separable CNN to extract key characteristics from images, paired with Softmax classifiers and Support Vector Machine (SVM) for precise categorization. The system was applied in a smart classroom to monitor student attendance, achieving a 98.11% success rate on the labeled faces in the Wild database. The cloud-based edge computing architecture demonstrated potential to enhance security and effectiveness in educational contexts, with notable improvements in accuracy and efficiency compared to existing methods.

Subject Classification: *68T07.*

Keywords: *FaceNet, CNN, Detection, LFW, SVM, Softmax.*

1. Introduction

Face recognition technology is gaining popularity for its speed and efficiency in identifying individuals across various sectors. It can simultaneously identify multiple people, making it a crucial component of secure authentication systems worldwide. Face recognition captures images from distance or extracts frames from video streams, processing them quickly without user interaction. This technology is particularly useful in high-security environments like airports, corporate buildings, and public events. Utilizing cutting-edge ML methods, including CNNs, has significantly improved the precision and reliability of facial recognition software. The use of large and diverse datasets has been instrumental in refining these technologies, enabling them to effectively analyze and learn from a vast array of visual data [1].

Face recognition technology is revolutionizing security and convenience in high-stakes environments by automating identification processes. It minimizes bottlenecks and enhances movement flow, making it indispensable in modern security infrastructure. As a result, face recognition technologies are becoming increasingly sophisticated and integral to numerous applications requiring reliable, rapid processing and identification capabilities [2].

The related work is covered in Section 2, and the proposed methodology—which includes a detailed description of a model using channel-wide separable dilated CNN and conventional methods—is covered in Section 3. The findings and discussion are presented in Section 4, and the conclusion and next steps are explained in Section 5.

2. Literature Review

Recent advancements in face recognition technology, particularly deep learning, have improved accuracy and operational efficiency. Schroff

et al. introduced FaceNet, which uses a mapping to a small Euclidean space to achieve advanced facial recognition capabilities. This system optimizes the embedding itself, setting a new accuracy record on the LFW dataset at 99.63% [3] developed a framework that combines two CNNs, LNet, and ANet, for effective face localization and high-level attribute recognition, improving attribute classification results and demonstrating the power of massive pre-training datasets. The study explored a deep learning method that incorporates explicit 3D face modeling, improving face recognition and achieving near-human performance on the LFW dataset [4]. Sun et al. proposed learning Deep-ID features through facial recognition exercises, enhancing face verification capabilities. Warman et al. applied face recognition to automate attendance systems, demonstrating the practical utility of combining deep learning models like FaceNet and ArcFace for efficient identity verification. Raspberry Pi was used for integrating face recognition technologies for lecture attendance management, achieving 100% accuracy in academic settings. The study highlights the practical utility of face recognition in real-world settings [5]. Seah developed a real-time attendance management system using face recognition technology, reducing human intervention and increasing safety. The system uses OpenCV and dlib libraries for face detection and recognition, and OneShot Learning optimizes data storage and processing in at-tendance systems [6]. The study highlights the benefits of fingerprint-based attendance systems in biometric security, comparing them to face recognition technologies [7]. It presents a non-intrusive method using a CNN model and K-Nearest Neighbor algorithm for facial recognition, achieving an attendance accuracy of 98% and showcasing its scalability for large educational institutions [8]. [9] compared three face recognition techniques, ANN, HMM, and PCA, found that ANN outperformed the others with a low error count and a recognition percentage of 97.49%, compared to 3.5% on the Yale Face database. A pilot project on Android integrates face recognition, biometric verification, and RFID systems [10].

The field of face recognition is advancing with deep learning techniques, improving biometric system accuracy and efficiency. These advancements enhance security, operational efficiency, and practical applications like automated attendance systems. Future research aims to integrate these technologies in dynamic environments for reliable face recognition systems.

3. Methodology

Initially, a facial dataset was acquired. This dataset underwent preprocessing, followed by face detection. Subsequently, feature extraction was performed on the detected faces, and the extracted features were classified. The classified results were then matched with other existing results, and the final outcomes were reported to the admin as described in Fig. 1.

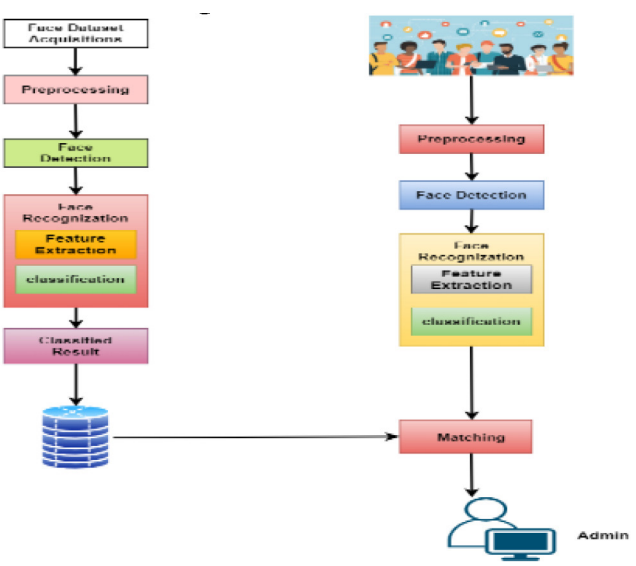


Figure 1

Proposed Methodology

3.1 *Traditional Method*

The DL-based method for image identification uses a deep Convolutional Neural Network for image properties and a SoftMax classifier for classification, enhancing performance [11]. The approach fills gaps in character predictions using improved SVM and Soft-Max classifiers, ignoring image pixel measurements. ReLU and Max Pooling are obtained on every convolution layer, and channel-wise and depth-wise convolution is achieved at the convolution level [12]. Convolutionally and pointwise convolution is a technique that expedites data processing and calculation by replacing the convolution block with a channel-wise convolution block, resulting in fewer calculations.

3.2 *Model with channel-wide separable dilated CNN*

The Channel-wise separability CNN that has been dilated combines traditional convolutional layers of simple CNNs with a more powerful, new technique that replaces these. This approach of the model ensures faster computation and reduced dimensionality compared with the previous models. There are fully connected nodes has three output channels, 64, 64, and 8 in each layer so overall, it is a global feature-based convolutional neural network explained in Table 1.

Table 1
Fully connected layer parameters

Attribute of the Layer	Fully Connected 1	Fully Connected 2	Fully Connected 3
Count of Activation Nodes	64 units of ReLU	64 units of ReLU	8 units of SoftMax

[p - k - (k - 1) * (d - 1) + 2 * maturity size] represents the convolution layer's output size in a dilated CNN./S + 1.

Here, d is the dilating rate, S is the [13] stride with respect to convolution, p is the padding, and k is the kernel or convolution filter size. Normal convolution using a 5 x 5 filter is twenty-five times more expensive than dilated CNN networks, which are more effective in terms of training time and accuracy. Compared to basic CNN, channel-wise separable convolution processes data more efficiently while having a lesser computational capacity. However, This model exhibits superior training accuracy compared to ordinary CNN or dilated [14]. The channel-wise separable convolution and dilated CNN model offers faster processing and improved training accuracy compared to previous CNN models. It can handle greater quantities of information in a given amount of time and requires less processing power than basic CNNs. When choosing step = +1, it is equivalent to choosing step size = store from the previous time step. Dilated CNN has low computing complexity and a channel-wise separable final output size dimension. Table 2 compares various CNN models using the SoftMax classifier. It reveals that channel separable convolution and dilated CNN are highly accurate during training, providing a good sense of performance points throughout the entire training process.

Table 2
Multiple CNN models with a SoftMax classifier

Type of Convolution	Iteration Count	No. of files used for training	Accuracy (%)
Simple CNN	128	5932	96.00
Dilated CNN	128	5932	97.00
Channel wise Separable CNN	128	5932	97.22
Channel wise Separable with Dilated CNN	128	5932	98.00

Additionally, Channel-wise separable CNN outperforms dilated CNN, while dilated CNN outperforms basic CNN in terms of performance [15]. The channel-wise separable with dilated CNN model outperforms the SVM classifier in terms of training accuracy and duration as shown in Table 3, while the SoftMax classifier requires fewer iterations but has greater training accuracy compared to the SVM classifier.

Table 3
Multiple CNN models with a SoftMax classifier

Type of Convolution	Iteration Count	No. of files used for training	Accuracy (%)
Simple CNN	128	5932	97.21
Dilated CNN	128	5932	98.01
Channel wise Separable CNN	128	5932	98.21
Channel wise Separable with Dilated CNN	128	5932	99.72

Results and Discussions

As illustrated in Fig. 2, in the institution above there are a great many classrooms. The data from individuals is transferred to the cloud, allowing for the registration of hazards and their aspects, eliminating the need for a throw rate or interval. Attendance tracking in intelligent classrooms allows for the identification of known attendees, with device names and edge node photographs being taken from the device, which can then be passed to a network interface. Diplography of the cluster of staff members and

students in the photo, i.e. Comes clearly (deadline) [16]. The cloud generates attendance reports based on predicted attendance times, sending photographs to a network interface. This algorithm allows an online application to check the operation of group face attendance systems. Students can add or modify their information, and staff can upload pictures of student groups. Groups can also add and modify attendance functions. The system is configured by an administrator for the user's username, preventing hacking and ensuring a secure environment for students.

Students and instructors use their registration numbers and professors' usernames to access a scholarship website. Users must update their Arcanum by filling out six pictures and personal data, which are encrypted and saved in a database. After registration, students can use face-based log-in methods and compare their images with previous ones. The scholarship office uses the database to track attendance and sends a list of present and absent scholars to the server admin. The office can then prepare a report for a day and submit it to management authorities. The proposed feature extraction system based on convolutional neural networks can be adapted for automatic image classification. It utilizes a channel-wise separable CNN to install the picture feature extractor and then utilizes either an SVM [17] or SoftMax to classify the pictures. A face recognition system for tracking student attendance on the Market-1501 Tagged Faces in the Wild dataset achieved 98.11 percent accuracy, with the SVM classifier showing better training accuracy than a SoftMax classifier, despite taking longer to train.

Table 4

Multiple CNN models with SoftMax Classifier

	Total Students	Detect Faces	Accuracy (%)	Recognize Faces	Accuracy (%)
Class 1	35	34	97.14	33	94.28
Class 2	31	30	96.77	28	90.32
Class 3	33	31	93.93	31	93.93
Class 4	26	26	100.00	25	96.15
Class 5	34	33	97.05	31	91.17

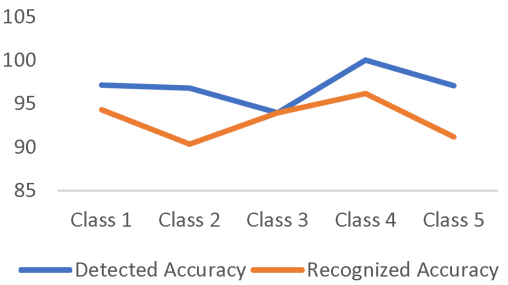


Figure 2

Accuracy Comparison between Detected and Recognized faces

During the workforce expended reported here, the Wild data groups. were used the dataset linked also to get over 13000 images from the network. Each image has a docket. As has been the case with faces to then analyze the expected algorithm quantitatively, the database contains not only both the name of the particular person and 5749 pictures of a wide variety of individuals. See Table 2, volunteers of the instructions Five got more or less face to the certificate. This database is drawn from many adept properties of face convolutional encoding. After the last layer of convolutional operation carries the convolution layer before it. According to Table 4, the desired mean accuracy for facial detection and identification is 96.9% and 93.17%, respectively, as shown in Figure 2. In deep learning approaches, the confirmed docket is compared to the forecast docket to verify the inaccuracy. After learning each example, one round of error, when validated with the training set, is substantially reduced. The error seems hard always to eliminate in the model form. Merely, in Mean Square Error(MSE) the measure by:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(\widehat{P}_n - P_n)^2 \tag{1}$$

N is data points; P'-P are forecasted and checked off the same sample; we will see how good is our planned model by testing it with data as well as what algorithms are relevant for studying on LFW wild dataset. Conclusions delivered by a predicted model are more excellent than those attained with the state-of-the-art method.

4. Conclusion and Future Direction

The study introduces a new convolutional neural network (CNN) architecture optimized for facial recognition and detection tasks. The

system extracts significant facial features using a channel-wise separable CNN, and improves classification skills with the integration of Support Vector Machine (SVM) and Softmax classifiers. When deployed in a smart classroom, it achieved a remarkable 98.11% accuracy on the LFW dataset, demonstrating its potential for reliable student attendance monitoring. The integration with an IoT-based edge computing framework offers real-time processing, efficient data management, and enhanced security.

References

[1] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 815–823. Accessed: Jul. 04 (2024).

[2] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep Learning Face Attributes in the Wild," presented at the Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 3730–3738. Accessed: Jul. 04 (2024). [Online].

[3] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701–1708. Accessed: Jul. 04 (2024). [Online].

[4] Y. Sun, X. Wang, and X. Tang, "Deep Learning Face Representation from Predicting 10,000 Classes," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1891–1898. Accessed: Jul. 04 (2024).

[5] G. P. Warman and G. P. Kusuma, "Face recognition for smart attendance system using deep learning," *Commun. Math. Biol. Neurosci.*, vol. 2023, no. 0, p. Article ID 19, Feb. (2023), doi: 10.28919/cmbn/7872.

[6] F. Azmi, A. Saleh, and A. Ridwan, "Smart Management Attendance System with Facial Recognition Using Computer Vision Techniques on the Raspberry Pi," *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 11, no. 1, pp. 38–44, Jan. (2023), doi: 10.55524/ijircst.2023.11.1.9.

[7] A. Lateef and M. Kamil, "Face Recognition-Based Automatic Attendance System in a Smart Classroom," *Iraqi J. Electr. Electron. Eng.*, vol. 20, no. 1, pp. 37–47, Jun. (2024), doi: 10.37917/ijeee.20.1.4.

[8] Y. Seah, "Smart attendance system with IoT-based face recognition using deep learning approach," other, UTAR, 2023. Accessed: Jul. 04 (2024). [Online]. Available: http://eprints.utar.edu.my/5821/

[9] A. Aggarwal, M. Alshehri, M. Kumar, P. Sharma, O. Alfarraj, and V. Deep, "Principal component analysis, hidden Markov model, and artificial neural network inspired techniques to recognize faces," Concurr. Comput. Pract. Exp., vol. 33, no. 9, p. e6157 (2021), doi: 10.1002/cpe.6157.

[10] E. Hajri, F. Hafeez, and A. V, "Fully Automated Classroom Attendance System," *Int. J. Interact. Mob. Technol.* IJIM, vol. 13, p. 95, Aug. (2019), doi: 10.3991/ijim.v13i08.10100.

[11] A. Lohar, S. Patel, R. Mehta, P. Shah, and N. Desai, "Smart Attendance System Through Facial Recognition," (May 2023). [Online]. Available: https://www.creativecommons.org/licenses/by/4.0/.

[12] M. Rahman, K. Rumman, R. Ahmmed, A. Rahman, and M. Sarker, "Fingerprint Based Biometric Attendance System," Eur. Chem. Bull., vol. 12, pp. 184–190, Mar. (2023), doi: 10.31838/ecb/2023.12.s3.026.

[13] R. Alqudah, A. Arafat, M. Shawish, and H. A. Shehadeh, "Biometric-based smart attendance management system using face recognition and authentication," AIP Conf. Proc., vol. 2979, no. 1, p. 020004, Oct. (2023), doi: 10.1063/5.0174784.

[14] N. A. M. Zulkefli, H. S. Kashoob, M. H. M. Adnan, and N. H. M. Ariffin, "Smart Attendance System Utilizing Face Recognition Technology Based on Principal Component Analysis," in Technology: Toward Business Sustainability, B. Alareeni and A. Hamdan, Eds., Cham: Springer Nature Switzerland, pp. 214–223 (2024), doi: 10.1007/978-3-031-54019-6_20.

[15] A. Khan and I. Sharma, "Machine Learning-Based Methodology for Preventing Ransomware Attacks on Healthcare Sector," in 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), pp. 1–5, No. (2023). doi: 10.1109/RMKMATE59243.2023.10368971.

[16] A. R. Singh, G. Singh, and N. Saluja, "Long-Short Term Memory Analysis of EEG Data Using Python for Emotion Classification," in 2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), pp. 1–5, Dec. (2023), doi: 10.1109/SMARTGENCON60755.2023.10442499.

[17] S. J. Karale, N. Mangrulkar, S. S. Badhiye, A. B. Patil, and D. Bhoyar, "Performance evaluation of machine learning algorithms for Parkinson's disease identification model," *J. Stat. Manag. Syst.*, vol. 27, no. 2, pp. 237–247 (2024), doi: 10.47974/JSMS-1250.

Received August, 2024