Auto-Attendance System using Face Detection and Recognition

Rohini Das(20BCB0098) | Dharmik Naicker(20BCB0148) | Shobhit Agrawal(20BDS0162) | Somit Jain(20BDS0181) | Prof. Rajkumar S| SCOPE

Introduction

- Recent advances in facial recognition technology have resulted from better models. Facial recognition is challenging due to the huge variety of faces in images, including position, expression, and lighting. Researchers advise using deep learning approaches, such as convolutional neural networks (CNNs), to extract high-level properties from pictures for face identification and recognition. This research focuses on face detection and identification using ResNet50, a deep CNN architecture that performs well in picture recognition. We use transfer learning to fine-tune the pre-trained ResNet50 model's weights and add new face-recognition layers. Our model is trained and evaluated using a dataset of face images in various postures, emotions, and lighting conditions. After preprocessing photos to recognize and extract faces, we split our dataset into training and validation sets. On the validation set, our ResNet50 model is evaluated after transfer learning on the training set.

Motivation

- Transfer learning with ResNet-50 can help picture categorization with low data. Fine-tuning the pre-trained model on your dataset yields great accuracy with less training data. ResNet-50, a deep convolutional neural network design, performs well on many image classification tasks.

SCOPE of the Project

- Our findings demonstrate that the improved ResNet50 model outperforms the original ResNet50 model in our face recognition test, delivering higher accuracy and fewer false positives. This highlights the efficiency of utilizing transfer learning to create a face detection and identification model based on ResNet50, which may have a variety of applications, such as security systems, surveillance systems, and digital marketing.

Methodology

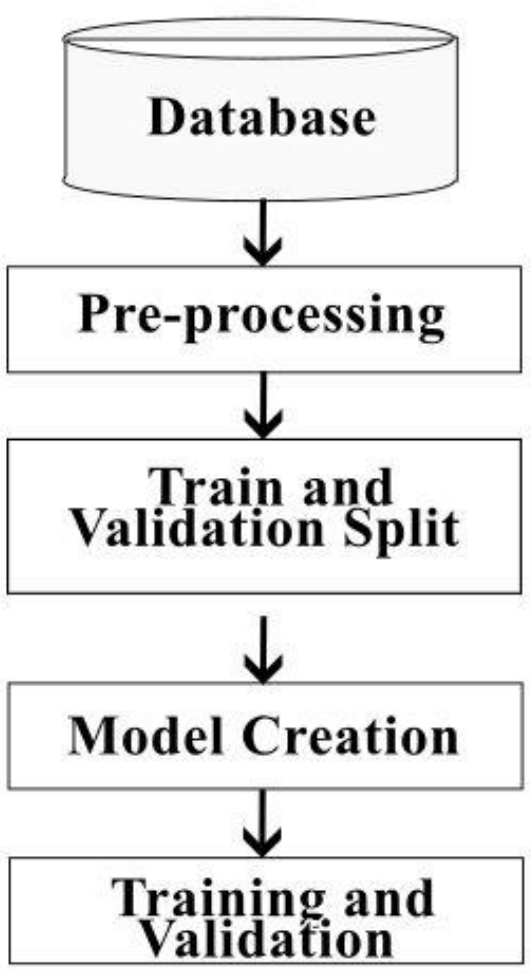

Fig 1. Block diagram of proposed model

1 Database: We have used the AT&T and FaceScrub database, which are openly accessible database with different images of distinct subjects taken at different time with varying lighting, facial expressions.

2 Preprocessing: After importing the dataset, we are resizing the images to 180*180 and changed the extension (.pgm to .png).

3 Train and Validation Split: The pre-processed data is now divided into train and validation sets, we split the dataset into 80-20.

4 Model Creation: ResNet-50, a convolutional neural network with 50 layers, is what we employ. A network that has been pretrained using more than a million photos from the ImageNet database was loaded. Then, we supplemented the pretrained model with our own layers

5 Training and Validation: After creating the model, we are running multiple epochs. Each epoch consists of training the model and validating it.

- We outline our proposed framework architecture for automatic attendance, in which we want to feature extract after pre-processing the input photographs, and then compare with already-existing images in the database to mark attendance. We employ transfer learning, a deep learning and machine learning technique that uses a previously trained model as a starting point for a different but related job. Transfer learning is fine-tuning a pre-trained model that has already learned to recognize specific features on a new dataset or task as opposed to building a model from start on a big dataset. Here, our pre-trained model is ResNet50. It is a design for a deep convolutional neural network, which is frequently used in computer vision tasks like object and picture detection. It uses skip connections, also known as residual connections, to solve the vanishing gradient problem that can arise in very deep neural networks. It has 50 layers. To ensure that our model does not overfit or underfit, we add a few additional layers and dropouts in addition to ResNet50.
- Our objective is to identify every face in the frame and concentrate on it individually. Even if a face is turned in a different direction or is present in poor lighting, features must be taken into account in order to identify it. The model should also be able to identify distinctive facial characteristics, such as the size of the eyes or the length of the face, that can be used to identify one individual from another. In order to ascertain the person's name, these extracted features will next be matched to the features of known people.

Result

Reference No.	Proposed Model	Existing Work	Proposed Model Accuracy	Existing Work Accuracy
[15]	ResNet-50 + 3 additional layers	ANN combined with modified MAML.	98.72%	96.20%
[16]	ResNet-50 + 3 additional layers	Residual CNN - Resnet 152v2 is the residual network variation used by the authors	98.72%	97%
[17]	ResNet-50 + 3 additional layers	CNN with the Approximation Wavelet Transformation	98.72%	97.42%
[18]	ResNet-50 + 3 additional layers	AlexNet-CNN architecture-based transfer learning framework	98.72%	fc6: 96% fc7: 98.17% fc8: 97%
[19]	ResNet-50 + 3 additional layers	AB-FR model, a convolutional neural network face recognition method based on BiLSTM and attention mechanism.	98.72%	96.79%

Table 1: Comparing Performance of Different Papers

- *Machine learning models may be evaluated by how well they perform on a validation dataset that was not utilised during training. It is frequently employed to approximate the model's adaptability to novel, unseen data.*
- *The validation set is used to check the model's accuracy and prevent overfitting while it is being trained during the machine learning process. The precision with which a model performs on the validation set is known as its validation accuracy.*

$$Validation\ Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

True positives (TP) - the cases where the model correctly predicts the positive class, False positives (FP) - the cases where the model predicts the positive class but the actual class is negative, True negatives (TN) - the cases where the model correctly predicts the negative class, False negatives (FN) - the cases where the model predicts the negative class but the actual class is positive.

Similar to accuracy, the range of validation accuracy as a metric is from 0 to 1. A validation accuracy of 0 indicates that the model is not able to correctly classify any of the cases in the validation set. A validation accuracy of 1 indicates that the model is able to correctly classify all the cases in the validation set.

Conclusion

In conclusion, the article proves that a transfer learning setup based on the ResNet-50 architecture with three extra layers can successfully complete a given job. The model was able to learn characteristics that were pertinent to the new dataset by drawing from the pre-trained ResNet-50 model on a big dataset. By including an additional three layers, the model was better able to adapt to the new challenge, leading to improved validation accuracy in comparison to previous works in the same field. Our suggested method makes use of many facets of picture classification, object recognition, and segmentation, and the high accuracy it achieves has the potential to boost the effectiveness and precision of automated systems across a wide range of fields, including the automation of classroom attendance. Overall, this research shows how transfer learning with pre-trained models, and the ResNet-50 architecture in particular, may help deep learning models do better on a variety of tasks. Our model may be fine-tuned for use with real-world data in a production setting with more study.

References

- Zhang, Ligang & Tjondronegoro, Dian & Chandran, V.. (2012). Discovering the Best Feature Extraction and Selection Algorithms for Spontaneous Facial Expression Recognition. Proceedings - IEEE International Conference on Multimedia and Expo. 1027-1032. 10.1109/ICME.2012.97.
- Viola, Paul & Jones, Michael. (2001). Robust Real-Time Object Detection. International Journal of Computer Vision - IJCV. 57.
- Yousaf, Muhammad Haroon & Baloch, Naveed Khan & Ahmad, Waqar & Baig, Muhammad. (2012). Algorithm for Efficient Attendance Management: Face Recognition based approach. international journal of computer science issues. 9.