IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 11, May 2023

To Study the Different Types of Face Recognization Algorithm

Dhanshree Adsul[1], Radhika Khadse[2], Harshada Khandre[3], Neha Khomane[4], Dr. Neelam Kumar[5]
Students, Department of Computer Engineering[1,2,3,4]
Professor[5], Department of Computer Engineering[5]
Shree Ramchandra College of Engineering, Pune, Maharashtra, India

Abstract: *The objective of this review paper is to examine various algorithms employed in face recognition. A face analyzer refers to software designed to authenticate or verify an individual's identity based on their facial features. It operates by identifying and measuring distinct characteristics of a face within an image. Facial recognition technology can detect human faces in images or videos, determine if two face images belong to the same person, or search for a specific face within a vast collection of images. Facial recognition is extensively used in biometric security systems to establish unique identification for user onboarding, and logins, and to enhance user authentication measures. Additionally, face analyzer technology is commonly integrated into mobile and personal devices for device security purposes.*

Keywords: Local Binary Patterns (LBP), Scale-Invariant Feature Transform (SIFT),Histogram of Oriented Gradients (HOG),Convolutional Neural Networks (CNN)

I. INTRODUCTION

Face recognition is an advanced biometric technology used for identifying or verifying individuals based on their facial features. It employs computer algorithms to analyze and match unique facial patterns, enabling automatic detection and recognition of individuals in images or videos.

Distinct facial features, including the arrangement of eyes, nose, mouth, and other facial characteristics, play a crucial role in recognizing and differentiating individuals. Face recognition technology emulates the human visual system by extracting and analyzing these facial features to make accurate identification or verification decisions.

The face recognition process generally involves three primary steps: face detection, feature extraction, and matching. Face detection algorithms are used to locate and isolate faces within an image or video frame. Once the faces are detected, feature extraction techniques analyze and extract specific facial characteristics, such as eye shape, distance between facial landmarks, or skin texture. These features are then transformed into a mathematical representation known as a face template or face signature.

The extracted face template is compared with a database of existing face templates to find a match. Matching involves measuring the similarity between the extracted features and the stored templates using algorithms like Euclidean distance, correlation, or machine learning-based classifiers. If a close enough match is found, the face is identified as belonging to a specific individual. In the case of verification, the extracted features are compared with a single template associated with the claimed identity to determine if they match.

Face recognition technology has made significant advancements, driven by progress in computer vision, machine learning, and deep learning algorithms. These advancements have improved the accuracy, speed, and scalability of face recognition systems, making them increasingly applicable in various domains.

Abbreviations and Acronyms

- Local Binary Patterns (LBP)
- Scale-Invariant Feature Transform (SIFT)
- Histogram of Oriented Gradients (HOG)
- Convolutional Neural Networks (CNN).

Local Binary Patterns (LBP)

Local Binary Patterns (LBP) is a widely utilized texture descriptor in the field of computer vision and has demonstrated its effectiveness in face recognition tasks. Initially introduced by Ojala et al. in 1994, LBP has gained popularity due to its simplicity and computational efficiency. This local texture descriptor captures the local structure and patterns within an image by comparing the intensity values of a central pixel with its neighboring pixels in a circular or rectangular neighborhood. The result of this comparison is encoded into a binary pattern, often represented as a binary number or a histogram.

The main steps involved in the LBP algorithm for face recognition are as follows:

1. Image Gridding: The face image is divided into a grid of cells, typically of equal size
2. Local Pattern Calculation: For each pixel in the image, the LBP operator is applied. This operator compares the intensity value of the central pixel with its surrounding neighbors. If the intensity of the neighbor is greater than or equal to the central pixel, it is assigned a value of 1; otherwise, it is assigned a value of 0. This comparison is performed for each neighbor in the neighborhood, resulting in a binary pattern.
3. Histogram Calculation: After computing the LBP values for all pixels in each cell, a histogram is generated by counting the occurrences of each LBP pattern within the cell.
4. Feature Extraction: The histograms from all the cells are concatenated to create a feature vector that represents the entire face image.

Advantages

- Robustness to Illumination Variations: LBP demonstrates robustness when faced with variations in lighting conditions. By capturing local texture patterns instead of relying solely on global intensity values, LBP is less affected by changes in illumination. This attribute makes LBP an appropriate choice for face recognition tasks that involve varying lighting conditions.
- Computational Efficiency: One of the notable advantages of LBP is its computational efficiency, as it demands minimal computational resources. The algorithm compares the intensity values of neighboring pixels and encodes the resulting patterns into binary codes. This efficiency makes LBP well-suited for real-time applications or situations where computational resources are limited.

Disadvantages

- Limited Discriminative Power: Although LBP excels at capturing local texture patterns, it may have limitations in capturing global facial structures or fine-grained details. Factors like pose variations or subtle differences between individuals that extend beyond local texture patterns may not be effectively captured by LBP.
- Sensitivity to Spatial Arrangement: The performance of LBP is influenced by the spatial arrangement of pixels within a neighborhood. Minor changes in this arrangement can lead to different binary patterns, which may impact the accuracy of recognition. This sensitivity can restrict its performance in scenarios involving significant facial pose variations.
- Lack of Rotation Invariance: LBP lacks inherent rotation invariance since the size and orientation of the circular or rectangular neighborhoods are fixed. Although extensions like Circular LBP partly address this limitation, recognizing faces under arbitrary rotations can still pose a challenge for LBP.
- Parameter Sensitivity: LBP relies on parameters such as neighborhood size and the number of neighbors, which require careful selection. Different parameter settings can yield varying results, making it crucial to choose optimal parameter values, which can be a challenging task.

Scale-Invariant Feature Transform (SIFT)

The Scale-Invariant Feature Transform (SIFT) algorithm is a widely utilized technique for feature extraction in computer vision. It was originally introduced by David Lowe in 1999 and has since become a fundamental tool in various tasks such as object recognition, image stitching, and 3D reconstruction. Key Point Detection: SIFT identifies

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 11, May 2023

distinctive key points or interest points in an image by detecting local maxima and minima in the scale-space representation. The scale-space representation involves generating a series of blurred images at different scales using Gaussian filters. Key points are identified as regions with significant responses in the difference-of-Gaussian (DoG) images.

Key Point Localization: SIFT refines the detected key points by eliminating poorly localized ones that have low contrast or are located on edges or corners. This refinement is achieved by comparing the intensity values and curvature of the scale-space extrema with their neighboring pixels.Scale and Rotation Invariance: SIFT achieves scale invariance by extracting key points at multiple scales through the scale-space representation. Key points are detected at various levels of image resolution, allowing SIFT to capture features at different sizes. Furthermore, SIFT computes an orientation for each key point based on the local image gradients, providing rotation invariance and enabling SIFT to robustly handle changes in the object's orientation.

Advantages

- Scale and Rotation Invariance: An important strength of SIFT lies in its capability to handle scale and rotation variations. By detecting and describing image features at multiple scales, SIFT can effectively match features even when objects or scenes change scale or rotation.
- Distinctive Feature Representation: SIFT excels at extracting highly distinctive features from images that are invariant to changes in illumination, noise, and minor geometric transformations. The feature descriptors generated by SIFT offer a dependable and concise representation of local image regions.

Disadvantages

- Computational Complexity: SIFT can exhibit high computational complexity, especially when dealing with large-scale datasets or real-time applications. The algorithm involves multiple stages, such as scale-space extrema detection, key point localization, orientation assignment, and feature descriptor computation. These stages collectively contribute to the computational overhead of the SIFT algorithm.
- Memory Consumption: SIFT generates a substantial number of feature descriptors, which can consume significant amounts of memory. This can pose limitations in memory-constrained environments or when processing large volumes of data.
- Sensitivity to Noise and Occlusion: Although SIFT is robust to many variations, it can be sensitive to noise and partial occlusions. Noisy or occluded regions may lead to unreliable or inconsistent key point detection and matching, affecting the overall performance of the algorithm.
- Parameter Selection: SIFT requires careful parameter selection, including configuring the scale space, setting key point thresholds, and defining descriptor parameters. Choosing appropriate parameter values is critical to achieving accurate and robust results. However, selecting optimal parameter values can be a challenging task, especially for users with limited experience in using the SIFT algorithm.

Histogram of Oriented Gradients (HOG)

The HOG (Histogram of Oriented Gradients) algorithm is a computer vision technique utilized for object detection and feature extraction. It was initially introduced by Navneet Dalal and Bill Triggs in 2005. HOG aims to capture and represent the local intensity gradients and orientations in an image to describe the underlying texture and shape information. The resulting HOG feature vector finds applications in various tasks such as pedestrian detection, human pose estimation, and object recognition. Traditionally, machine learning algorithms like Support Vector Machines (SVM) or classifiers are trained on labeled HOG features to perform object detection or classification. While HOG has been widely adopted and proven effective in certain contexts, recent advancements in deep learning, such as Convolutional Neural Networks (CNNs), have surpassed its performance in many computer vision tasks. These deep learning approaches automatically learn and extract more complex and discriminative features directly from the data, outperforming HOG in various scenarios.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 11, May 2023

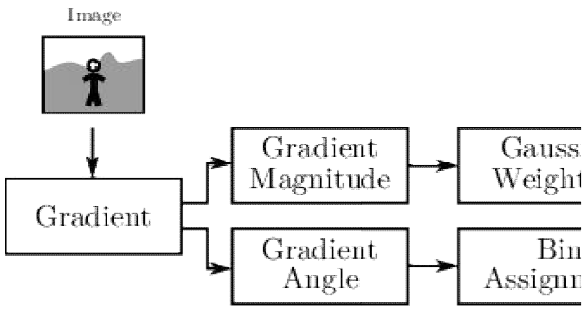

Fig .1 Architecture of HOG

Advantages

- Robust to Illumination Changes: The HOG features are derived from local gradient information instead of absolute pixel values, rendering them less susceptible to variations in lighting conditions. This robustness to illumination changes makes HOG effective in object detection tasks across different lighting environments.
- Translation and Scale Invariance: HOG exhibits the ability to handle object detection at various scales and locations by utilizing a sliding window approach and incorporating multi-scale analysis. This capability enables HOG to detect objects of different sizes and positions within an image.
- Descriptive Power: HOG captures significant shape and edge information in an image, making it highly effective in object recognition tasks. The orientation histograms offer a concise representation of local image structures, enabling robust discrimination between different objects.

Disadvantages

- Sensitivity to Occlusions: Since HOG relies on local gradient information, it can be sensitive to occlusions or instances where objects overlap. Occlusions can disrupt the gradients, leading to decreased accuracy in object detection and recognition
- Limited Rotation Invariance: HOG does not possess inherent capabilities for handling rotation variations. Different orientations of an object are treated as separate features. To achieve rotation invariance, additional techniques such as introducing overlapping blocks or incorporating orientation compensation may be required.
- Lack of Fine-grained Texture Details: While HOG excels at capturing global shape and edge information, it may not be as effective in capturing fine-grained texture details of objects. This limitation can impact the recognition of objects with texture-dependent characteristics.
- Parameter Sensitivity: The performance of HOG can be sensitive to parameter choices, including the number of histogram bins, cell size, and block size. Selecting appropriate parameter values is crucial to ensure accurate and robust results when using HOG.

Convolutional Neural Networks (CNN).

Convolutional Neural Networks (CNNs) are a type of deep learning algorithm specifically designed for processing structured grid data, such as images. CNNs have significantly impacted the field of computer vision and achieved state-of-the-art performance in various tasks, such as image classification, object detection, and image segmentation. CNNs consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers. Convolutional layers play a crucial role by applying learnable filters, also known as kernels, to the input data. These layers perform convolution operations to extract spatial features from the input images.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 11, May 2023

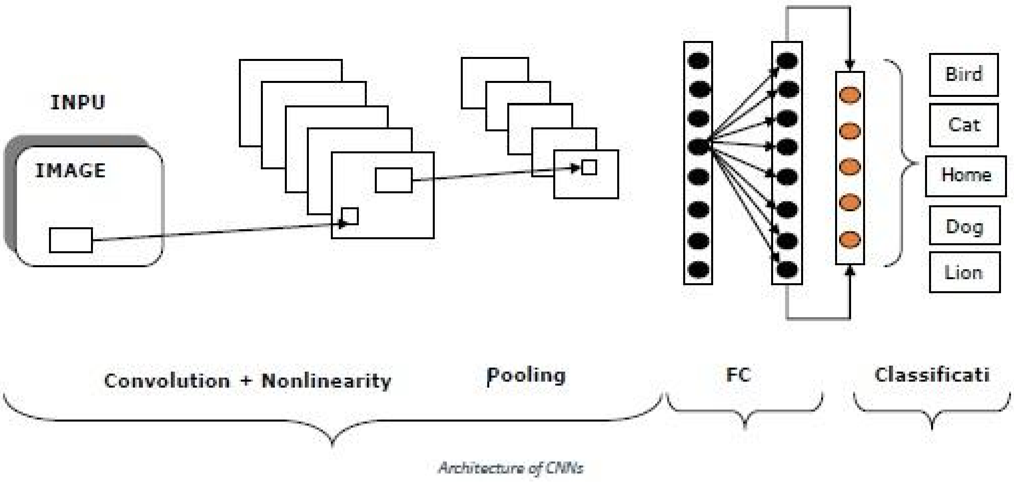

Fig.2 CNN Architecture

One of the key advantages of CNNs is their ability to automatically learn hierarchical representations of the input data. Through the convolutional layers, CNNs capture increasingly complex and abstract features, enabling effective representation of visual patterns in the input images. Pooling layers are commonly used after convolutional layers to down-sample the feature maps and reduce their spatial dimensions. This process retains important information while discarding unnecessary fine-grained details. Non-linear activation functions, such as ReLU, introduce non-linearity into the network and facilitate the learning of complex relationships between the input data and the desired output.

CNNs are typically trained using the backpropagation algorithm, which adjusts the network's weights based on the error between the predicted output and the ground truth labels. This optimization process minimizes a loss function, such as cross-entropy, through iterative weight updates using gradient descent. CNNs can also benefit from transfer learning, where pre-trained models on large-scale datasets like ImageNet are utilized as a starting point for new tasks. Transfer learning leverages the learned features from pre-trained models, leading to improved performance with smaller labeled datasets and reduced training time.

Spatial invariance is a fundamental property of CNNs. By exploiting this property, CNNs share learned features across different spatial locations, enabling the recognition of objects or patterns regardless of their position in the input image. This spatial invariance makes CNNs robust to translation and local spatial variations.

CNNs have found successful applications in various computer vision tasks, including image classification, object detection, semantic segmentation, facial recognition, and image generation. They have achieved remarkable performance, outperforming traditional methods in numerous benchmark datasets and real-world applications

Advantages

- Hierarchical Feature Learning: One of the key advantages of CNNs is their ability to automatically learn hierarchical representations of input data. They can capture increasingly complex and abstract features at different layers without the need for explicit feature engineering. This capability enables CNNs to extract meaningful features directly from raw data.
- Spatial Invariance: CNNs take advantage of the spatial invariance property of images, allowing them to recognize objects or patterns irrespective of their position in the input image. This property makes CNNs robust to translation and local spatial variations, making them well-suited for tasks like object detection and image classification.
- Parameter Sharing: CNNs employ parameter sharing across different spatial locations, reducing the number of learnable parameters. This sharing mechanism enables CNNs to efficiently capture spatial patterns and generalize well to unseen data, even with limited training examples.

- Local Receptive Fields: Convolutional layers in CNNs have local receptive fields, which means they only consider a small neighborhood of the input data at a time. This local perspective enables CNNs to capture local dependencies and spatial structures, enabling effective modeling of patterns in images.
- Transfer Learning: CNNs can leverage pre-trained models on large-scale datasets and transfer the learned features to new tasks with smaller labeled datasets. Transfer learning facilitates faster training and improved performance, particularly in scenarios where labeled data is scarce or limited.
- Parallel Processing: CNNs can be efficiently trained and executed on parallel hardware, such as GPUs (Graphics Processing Units). This parallel processing capability enables faster training and inference times, making CNNs well-suited for real-time applications and computationally demanding tasks.

Disadvantages

- Large Training Data Requirements: CNNs often demand a significant amount of labeled training data to achieve optimal performance. In certain cases, it may be challenging to obtain sufficient datasets, especially in niche or specialized domains. This limitation can hinder the effectiveness of CNNs when dealing with limited training data.
- Computational Complexity: CNNs, particularly deep architectures with numerous layers, can impose a substantial computational burden during training and require considerable computational resources. The training process typically involves multiple iterations and can be time-consuming, particularly when performed on CPU-based systems.

II. CONCLUSION

In this study, we examine various face recognition algorithms, including CNN, HOG, LBP, and SIFT, among others. Based on our analysis, we find that CNN is particularly well-suited for face recognition. CNNs possess unique characteristics that enable them to automatically learn and extract hierarchical features from images or grid-like data. Through the utilization of convolutional layers and pooling layers, they can effectively capture local patterns and spatial relationships, enabling robust analysis of visual data. As a result, CNNs have achieved state-of-the-art performance in a range of computer vision tasks. In contrast, HOG relies on handcrafted features based on gradients and histograms, which may not effectively capture complex patterns or image variations.CNNs have consistently demonstrated superior performance in tasks such as image classification, object detection, and image segmentation, primarily due to their capability to learn intricate features directly from the data.

REFERENCES

[1] Shelke DR Neelam Labhade Sujata Khedkar,Pratiksha Dhumal," Facial Recognition With Convolutional Neural Network For Driver", GIS SCIENCE JOURNAL, Volume9, ISSUE 5, Pages 2063-
2069H

[2] J. Shi, X. Zhu, and N. Niu, "A Research on Multi-Layer Perceptron Diagnosis Model Based on D Matrix", 2018 Prognostics and System Health Management Conference (PHM-Chongqing), pp. 769- 773, 2018.

[3] Neelam Labhade-Kumar, Dr.Yogesh Kumar Sharma, Dr. Parul Arora "Key Feature Extraction for Video Shot Boundary Detection using CNN", International Journal of RecentTechnology and Engineering (IJRTE) blue Eye (Scopus), ISSN: 2277-3878, Volume-8 Issue-5, January2020RetrievalNumber:E6789018520/2020© BEIESPDOI:10.35940/jute.E6789.018520,PP4763-4769

[4] P. J. Phillips and A. J. O'Toole, "Comparison of human and computer performance across face recognition experiments", Image and Vision Computing, vol. 32, no. 1, pp. 74-85, 2014.

[5] Z. B. Lahaw, D. Essaidani, and H. Seddik, "Robust Face Recognition Approaches Using PCA ICA LDA Based on DWT and SVM Algorithms", 2018 41st International Conference on Telecommunications and SignalProcessing (TSP), pp. 1-5, 2018.

[6] N. Sabri et al., "A Comparison of Face Detection Classifier using Facial Geometry Distance Measure", 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC), pp. 116-120, 2018.

[7] T. F. Cootes, D. Cooper, C. J. Taylor, and J. Graham, "Active shape models—their training and application", CVIU, vol. 61, no. 1, pp. 38-59, 2004.
[8] J. M. S. Belongie and J. Puzicha. Shape matching object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 24(24):509-522, 2002.
[9] S. Belongie, J. Malik, and J. Puzicha. Matching shapes. The 8th ICCV, Vancouver, Canada, pages 454–461, 2001.
[10] V. de Poortere, J. Cant, B. Van den Bosch, J. de Prins, F. Fransens, and L. Van Gool. Efficient pedestrian detection: a test case for SVM based categorization. Workshop on Cognitive,Vision2002.