Enhanced Face Detection in Videos Based on Integrating Spatial Features (LBP, CS-LBP) with CNN Technique

Faqeda Hassen Kareem[1,*], Mohammed Abdullah Naser[1]
[1] College of Science for Women, University of Babylon, Iraq
Emails: faqeda.albermany.gsci141@student.uobabylon.edu.iq; wsci.mohammed.abud@uobabylon.edu.iq

Abstract

Face detection is a crucial aspect of computer vision and image processing, in order to enable the automatic detection and identification of human faces in video streams, face detection is an essential component of computer vision and image processing. Applications for facial recognition, video analytics, security systems, and surveillance all depend on it. Face identification techniques face many obstacles and issues, such as positional fluctuations, illumination changes, resolution and scale issues, facial emotions, and cosmetics. Robust algorithms are required for efficient face detection. This field looks at the feature extraction process using a variety of techniques. These consist of the center symmetric local binary patterns (CS_LBP) approach and the local binary patterns (LBP) method. The YouTube Face database provided the video frames that we used for our study. In order to train the convolutional neural network (CNN) to detect human faces in the video and draw a bounding box around them. The experimental results of the suggested approaches show that. The accuracy rate was 94% higher with the LBP techniques. However, the CS_LBP technique showed the best level of accuracy in both face detection and face rectangle recognition, with an accuracy rate of 95%.

Keywords: Face Detection; Spatial Feature Extraction; CNN; LBP; CS_LBP

1. Introduction

Face detection is a significant and crucial aspect of computer vision systems, focused on extracting information from facial images. Face detection is a crucial initial stage in face verification Facial recognition and clustering, facial landmarks, facial hallmark categorization ,and face tracking [1]. Several face-identification techniques have been developed in the last decade. Despite significant progress in the discipline over the past few decades, accurate and effective face detection in natural environments still needs to be improved. Facial recognition relies on position, occlusion, scale, illumination, image quality, facial expressions, and other features. Face detection differs from standard object detection due to its lower component ratio variations and significantly bigger scale adjustments, ranging from several pixels to thousands of pixels [2]. The traditional approach involved extracting engineered features from the image and utilizing multiple classifiers to accurately identify facial areas, serving as the basis for first face detection endeavors. The Haar cascade classifier [3] and Histogram of Oriented Gradients (HOG) with Support Vector Machine (SVM) [4] are two significant classical works in face detection. This research exemplifies the latest cutting-edge accomplishments.

However, the precision of face detection in challenging photos with unresolved changes remains restricted in the WIDER FACE facial detection dataset[5]. Deep convolutional neural networks (CNN) have notably succeeded in various computer vision tasks such as image classification, object detection, and semantic segmentation [6]. These deep learning algorithms can bypass manual design processes and can be evaluated using established benchmarks like the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), which differs from conventional computer vision systems [7]. There is a growing research interest in using deep learning to deal with face

DOI: https://doi.org/10.54216/JCIM.140225
Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024

identification difficulties due to the increasing importance of deep learning in computer vision. Object detection has advanced significantly [8], drawing on established techniques and incorporating the latest developments in deep learning. Face detectors can yield superior detection results compared to traditional cascaded classifiers such as Faster R-CNN [8], YOLO[9], or single-shot detector (SSD)[10]. There are several similar works, such as Face R-CNN and Face R-FCN [11], that have been enhanced and modified using R-FCN and faster R-CNN [7]. Moreover, additional detectors include multitasking cascaded convolutional networks (MTCNN) [2].

The Main contribution of this research is the use of the LBP and CS_LBP algorithms to extract features from video frames and then input images of features into CNN for efficient face detection. The goal was to increase detection accuracy by extracting significant features from the image, thereby improving the CNN's performance. The following sections of this work are organized subsequently: Section 2 presents a comprehensive summary of the prior research carried out on face detection. Section 3 offers a thorough explanation of the proposed approach for detecting faces. Section 4 contains the presentation of the experimental result and discussion. In contrast, Section 5 acts as the conclusion of the research and provides an analysis of the future directions of the suggested approach.

2. Related Works

This section offers a comprehensive review of several procedures used for face detection, ranging from early techniques to current breakthroughs. Recently, numerous approaches have been extended to recognize and identify facial features.

Ning Zhang et al. (2020) [2]. They present a comprehensive analysis of Two-stage and One-stage detection models for face detection problems. Furthermore, it introduces a multi-task convolutional neural network (MTCNN) that attains an accuracy of 85.7% when utilized on images of wider faces. The system employs the MTCNN face detection model to accurately detect and identify the faces and their associated facial landmarks inside the video frame.

Aniruddha Srinivas Joshi, et al (2020)[12]. The system employs the MTCNN face detection model to accurately detect and recognize the faces and their relevant facial landmarks inside the video frame. A contemporary classifier is employed to examine facial photos and indicators, utilizing the MobileNetV2 architecture as an object detector to recognize regions that are concealed by masks. The proposed architecture was evaluated using a dataset comprising films that capture the motion of individuals in public areas while adhering to COVID-19 safety measures. The efficacy of this system in identifying facial masks was proved by reaching a precision rate of 81.74% on the selected dataset.

Mliki , et al (2020) [11]. The researcher presented a technique for enhancing the architecture of the Region-based CNN (Faster R-CNN) by incorporating both global and local information into an architectural framework. The design has two main networks: the region proposal network and the second network, which differentiates between facial and non-facial entities. The suggested approach achieves a recall rate of 92.09% based on experimental results.

Chen ,and Weijun (2021) [13]. This study intends to utilize YOLO-face, a face detector based on YOLOv3. YOLO-face incorporates a regression loss function and anchor boxes, enhancing face identification accuracy. The revised technique maintains fast detection speed and enhances precision, outperforming YOLO and its variations in evaluations conducted on WIDER FACE and FDDB datasets. The strategy achieved 78%, 73%, and 47% accuracy rates for easy, medium, and complicated photos, respectively.

Nandkumar Kulkarni et al. (2022)[14] . The Haar Cascade Classifier (HCC) is well-suited for real-time face identification, as demonstrated in this research. The HCC algorithm exhibits a face detection rate exceeding 90% when applied to pictures with a simple background. In addition, it attains a face detection accuracy of 93.24% when presented with images that contain a complex background.

Yilin Liu, and et al (2022) [15]. The researcher provides evidence for the validity of the prediction produced by target tracking and establishes the connection between face identification and target tracking. Incorporates the tracking results into the tracking algorithm as a temporal feature. Rest-SSD, a Single Shot Multi-Box Detector, decreases the probability of the occurrence of the incident by detecting faces. The issue of missing detection occurs due to alterations in facial posture and occlusion. The experimental results indicate that the simple, medium, and complicated images achieved accuracy rates of 93%, 92%, and 83%, respectively.

DOI: https://doi.org/10.54216/JCIM.140225
Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024

3. Proposed Face Detection Method

The suggested model detects faces in the video by converting the video into frames, pre-processing the frames. In a subsequent phase of the model, we employ one of the above strategies to extract the features and input them into CNN to evaluate the detection accuracy associated with each method.as depicted in Figure 1.

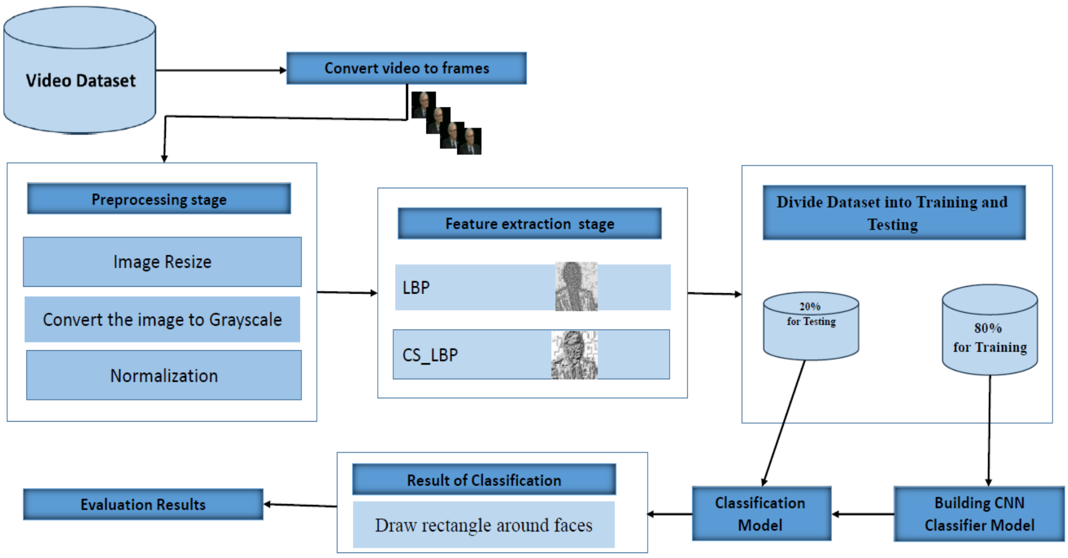


Figure. 1. Block diagram of the proposed system

3.1. The Dataset Collection: The YouTube Faces dataset[16][17] is used, often known as YTF, consists of over 3,000 face films showcasing more than 1,000 individuals that were obtained from YouTube. On average, there are two videos per person. The video clips have different durations, ranging from 46 frames for the shortest clip to approximately 6,000 frames for the longest one. The collection contains video clips with an average duration of approximately 180 frames.

3.2. Data Pre-processing: It is a systematic application of actions to an image or frame, preparing it for subsequent stages, facilitating data arrangement and filtering, and streamlining the dataset for feature extraction. The preprocessing stage involves implementing the following processes on the dataset:

3.2.1. Grayscale Image Conversion: To reduce the number of channels in the input image, converting these images (frames) from three-dimensional RGB bands to grayscale is necessary. Utilizing grayscale images enhances the speed and efficiency of the verification process.

3.2.2. Image Resizing: The process of resizing images can be incorporated into preprocessing techniques to standardize the image's dimensions or decrease computational complexity.

3.2.3 Normalization: CNN collects and analyzes a diverse array of images ranging from 0 to 1. In order to get the intended result, every pixel having a value ranging from 0 to 255 is rescaled by dividing it by 255, resulting in a new value between 0 and 1.

3.3. Spatial Feature Extraction: The spatial feature extraction process retains significant importance in various applications such as diagnosis, classification, clustering, recognition, and detection. The process entails converting unprocessed data into features that effectively capture fundamental aspects of the data[18]. The selection of features has a direct impact on the performance and correctness of an application. Feature engineering necessitates specialized knowledge and comprehension of the data. To enhance performance, researchers engage in the refinement and improvement of feature sets. The methodology extraction is contingent upon the characteristics of the data, the available computational resources, and the specific use cases[18].

Block-Level Feature Extraction: Block-level feature extraction from images involves analyzing an image by partitioning it into smaller units called "blocks" and extracting distinct features from each block separately. This technique aims to comprehend the image's framework and extract significant data.

- **Local Binary Pattern (LBP):**

Ojala et al [23]. were the first to introduce it. This methodology, which relies on statistical analysis, is a powerful method for determining textures and showcases its utility in this context. Multiple versions of LBP are frequently employed in face analysis due to their high classification performance. The LBP algorithm optimizes mutual

DOI: https://doi.org/10.54216/JCIM.140225
Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024

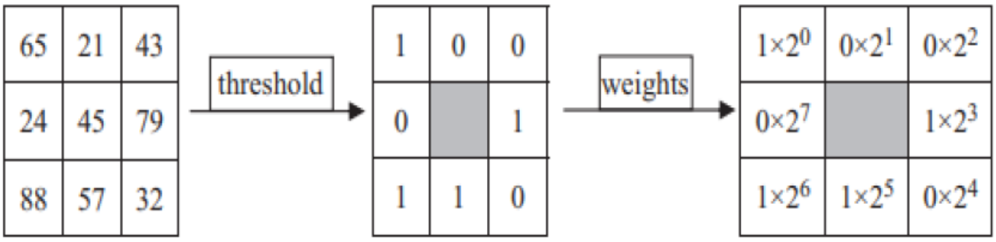
information and performs better in face analysis [24]. The calculation of LBP includes the comparison of the intensity of the grey level in the central pixel with that of its adjacent pixels. The binary value of 1 or 0 is assigned to each adjacent pixel, depending on whether it is lighted more or less than the central pixel. Following this, the binary entities are combined to produce a binary sequence that can be utilized as an image attribute[19]. The LBP result can be represented by equations (5) and (6)[20].

LBP p, r (X c ,Y c)=$\sum_{n}^{7} = 0\, \delta(g_n - g_c)$ (5)

Where gc = center pixel value positioned at (xc, yc), gn = one of the eight surrounding center pixel values with the radius R, P = is the whole neighborhood number, and a sign function δ(.) is defined such that

$$\delta(x)=\begin{cases}1 & if\ x \geq 0 \\ 0 & otherwise\end{cases} \quad (6)$$

Figure 2 depicts a direct implementation of the Local Binary Patterns (LBP) technique in the following manner.



*LBP code of the center pixel **45**: 1×20+0×21+0×22+1×23+0×24+1×25+1×26+0×27=1+8+32+64=**105**

Figure.2. An example of the LBP [21].

3.3.3. Center-Symmetric Local Binary Pattern (CS-LBP):

It signifies an additional altered iteration of LBP. The initial proposal aims to mitigate certain limitations of the conventional LBP in reducing dimensions. Consequently, the algorithm's computational complexity is significantly decreased[22]. A decrease in dimension results in reduced computational time, and processing speed is a significant metric for operators. Additionally, CSL_BP exhibits enhanced anti-noise capabilities by assuming that an image edge traverses a specific pixel.

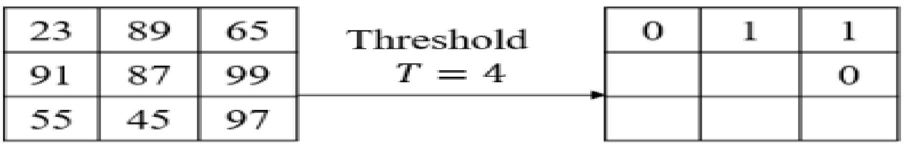
In the context of CS-LBP, the comparison is made between the center symmetric pairs of pixels, as opposed to the grey levels of individual pixels and the center pixel. The gradient operator exhibits a strong correlation with it. It considers the differences in grayscale between neighboring pairs of pixels in a given area. Hence, the CS-LBP approach incorporates both gradient-based features and LBP techniques. The CS-LBP characteristics can be calculated using equations (7) and (8)[23] .

CS-LBp,r,t=$\sum_{i=0}^{\frac{n}{2}-1} s(|gi - gi + n/2|)\, 2^i$ (7)

$$S(x)= \begin{cases}1 & if\ x \geq t \\ 0 & otherwise\end{cases} \quad (8)$$

Where gi and $gi{+}N2/$ equate to the grayscale of N centrally symmetric pairs of pixels distributed evenly around a circle of radius r. And N is the number of pixels other than the center pixel; for a 3 × 3 region, for instance, N = 8. T is a less-than-positive number that is used to increase the CSLBP operator's flexibility while smoothing the difference between grayscale images.

In the Figure (3), the threshold T is set to 4, and then the gray difference of four pairs of center symmetric pixels in the 3×3 window area is compared. When the threshold is less than T, the corresponding position will be set to 0, otherwise 1 will be set, and the CSLBP code value will be 0110.



CS-LBP code of the center pixel **87**:0×20+1×21+1×22+0×23=2+4=**6**

Figure.3. An example of the CS-LBP

DOI: https://doi.org/10.54216/JCIM.140225
Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024

The CS-LBP operator extracts a feature dimension of 16, significantly smaller than the feature dimension obtained by the LBP operator. This reduction in feature dimension is achieved during the statistical histogram technique, effectively achieving the reduction objective. Data is acquired linearly and subsequently stored and processed[24].

3.4. Convolutional Neural Network (CNN):

CNN is one of the most often used multi-layered deep neural network types in computer vision applications. It is a helpful tool for image recognition, object detection, and visualization. In AI, CNNs are the most popular neural networks for Deep Learning and image processing because they can handle large amounts of data, do not require costly segmentation, and do not require manual feature extraction [25].

The CNN model is built from these layers (Convolution layers, Max pooling layers, and Fully Connected Layer), According to Figure 4, the CNN architecture is depicted.

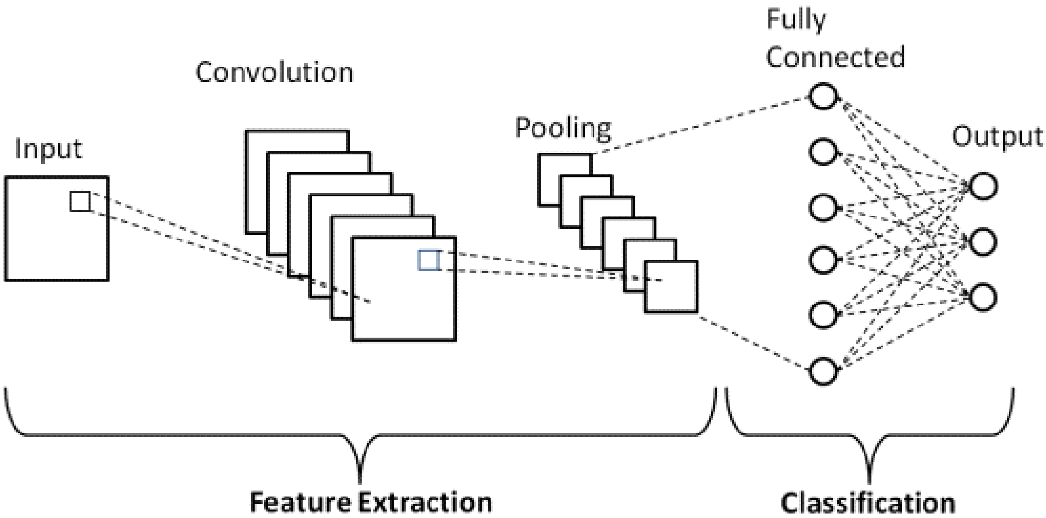


Figure.4. The architecture of convolutional neural network[26]

- **Convolution layer** is a crucial element in CNN, extracting features from an input image using mathematical techniques. Convolution filters generate feature maps using input data arrays. The output is gathered for each location point, resulting in a feature map[27].
- **The max pooling layer** decreases the number of parameters in the final model and mitigates overfitting by lowering the dimensions of the feature map. Typically, it is positioned with a down sampling ratio of 2:1 following the convolutional layer. Down sampling reduces the size of a map feature while preserving its depth, hence decreasing image noise. There are two primary forms of pooling: average pooling and maximal pooling. Max-pooling is a process that chooses the maximum value within a specific window, whereas average pooling calculates the average value from the input pixels. [26].

Flattening is a process that transforms two-dimensional arrays into a unified linear vector.

- **Fully connected layer** uses a pixel from the flattened matrix as input to classify an image. Fully connected layers (FC) are incorporated into the CNN design to complete the process. The FC layer connects nodes obtained after flattening, classifying input images into multiple groups based on training data. Dense is used in the convolutional neural network construction process. The final layer includes a loss function to optimize the model and improve prediction accuracy [29].

3.5. Evolution result:

Accuracy: refers to the ratio of accurately predicted observations to the total number of observations. Ratio of true positive predictions to the overall number of positive predictions[6].

$$\text{Accuracy}=\frac{TP+TN}{TP+FP+FN+TN} \qquad (6)$$

DOI: https://doi.org/10.54216/JCIM.140225
Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024

4. Result and Discussion

The proposed system being evaluated completed testing on a PC with a Core i7 processor and 16GB of RAM. The training was performed on the YouTube Face dataset, which consisted of 50 specifically chosen videos, totaling around 12,441 frames. The adopted dataset was divided into 80% for training and 20% for testing purposes. The frames were thus converted to grayscale images as a step in the pre-processing phase.

In general, using feature extraction approaches significantly enhances detection accuracy. The statement emphasizes the significance of utilizing specialized knowledge in a particular field and combining different approaches to improve the skills of deep learning models. This is particularly important when working with intricate, distorted data, such as low-resolution video frames. Based on the experimental results and figures (5,6), the following observations can be made:

Based on the given results, the following observations can be made:

1. LBP + CNN (Figure 5): Applying LBP feature extraction to video frames before utilizing CNN yielded a detection accuracy of 94%. LBP focuses on capturing local texture patterns. The LBP features likely provided CNN with more discriminative information for face detection, contributing to the improved accuracy.

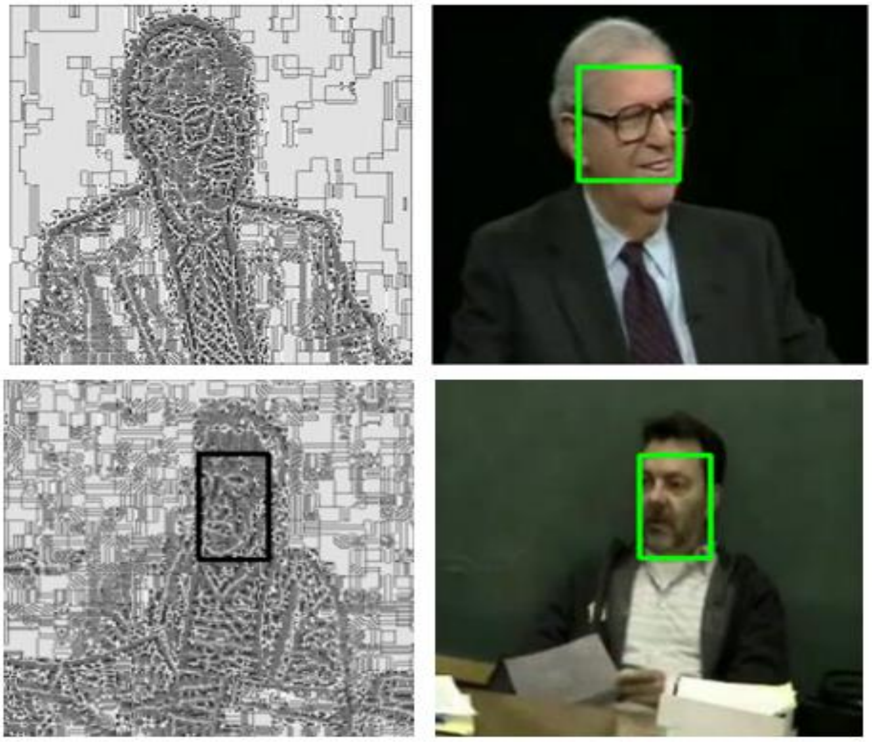


Figure.5. Face Detection with LBP Method

2. CS_LBP + CNN (Figure 6): Before applying CNN, using CS_LBP for feature extraction resulted in the highest accuracy of 95%. CS_LBP extends LBP by considering spatial relationships and symmetry. By incorporating spatial information and symmetrical patterns, CS_LBP better captures local texture patterns, leading to improved detection performance.

DOI: https://doi.org/10.54216/JCIM.140225
Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024

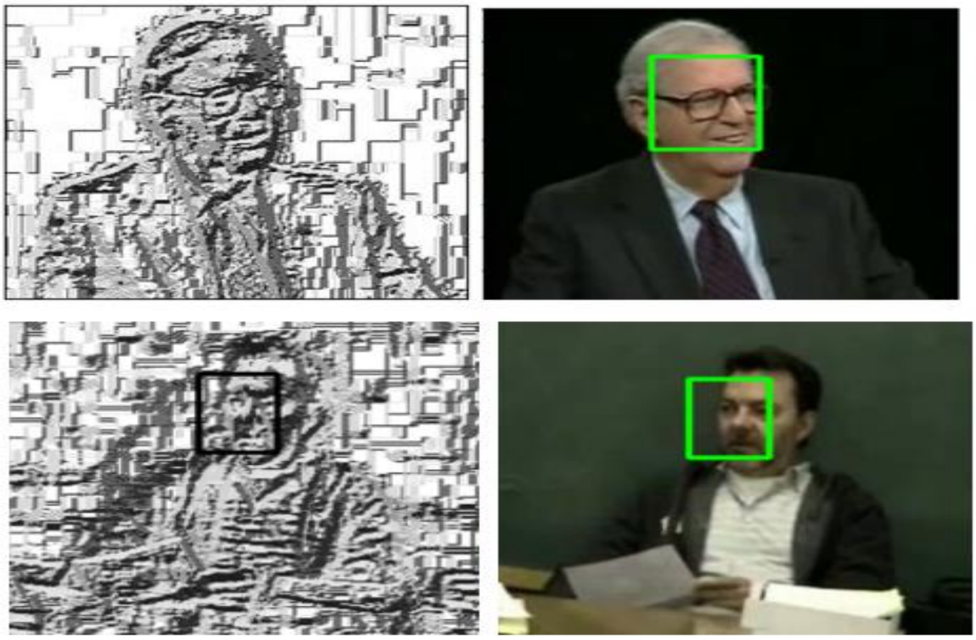


Figure.6. Face Detection with CS_LBP Method

Regarding accuracy, the CS_LBP + CNN method (Figure 6) is the best among the tested approaches, achieving the highest accuracy of 95%. It outperforms the LBP + CNN by providing a more robust representation of local texture patterns while considering spatial relationships and symmetry. The inclusion of spatial information and the consideration of symmetry likely contribute to better handling of variations in pose, illumination changes, occlusion, and other challenges faced in face detection.

Overall, the CS_LBP + CNN method demonstrates the best accuracy, indicating its effectiveness in improving face detection performance.

After evaluating the suggested approaches, we discovered that CS_LBP achieved the maximum accuracy in both face detection and identifying a face rectangle. With CS_LBP, it achieved an accuracy rate of 95%, surpassing LBP. Therefore, we will select CS_LBP with CNN as the proposed system.

The table (1) presents a comparison of different machine learning and deep face detection methodologies. The proposed method surpassed the previous researchers' work in terms of classification metrics, specifically accuracy, while utilizing a different dataset.

Table 1: Comparison with Related Works

Researcher(s) Name	The Year	technique (s)	Dataset	Accuracy
Ning Zhang et al[2].	2020	MTCNN	wider face data set, Pascal VOC database	85.7%
Aniruddha Srinivas Joshi, et al [12].	2020	MTCNN	chosen dataset	81.74%
Mliki, et al [11].	2020	Faster R-CNN	WIDER FACE and FDDB datasets	92.09%
Chen, and Weijun [14].	2021	YOLO-face	WIDER FACE and FDDB datasets	66%*
Nandkumar Kulkarni et al [15].	2022	HCC	Collected data from cameras	91%*
Yilin Liu, and et al [16].	2022	Rest-SSD	WIDER FACE and FDDB datasets	89%*
Our proposed	**2024**	**LBP &CNN**	**YouTube Faces dataset**	**92%**
		CS_LBP&CNN		**94%**

***** The average accuracy for images classified as easy, medium, and hard

DOI: https://doi.org/10.54216/JCIM.140225
Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024

5. Conclusions

Utilizing CNNs to extract features from video frames is effective, particularly when working with low-resolution and ambiguous data. Combining methods such as (LBP and CS-LBP) with CNNs can effectively enhance the accuracy of face detection tasks.

LBP and CS-LBP are conventional techniques used to extract features from images. These approaches are designed to capture many elements of image structure. When CNNs are paired with these approaches, which excel at automatically learning hierarchical features, the model's capability to accurately distinguish faces can be improved, even in difficult conditions. The empirical findings from the suggested approaches demonstrate that the accuracy rate of the LBP with the CNN model was 94% higher. Nevertheless, the CS_LBP with CNN model had the best level of accuracy, reaching 95%, in both face detection and face rectangle recognition. The superiority of CS-LBP can be attributed to its inclusion of spatial relationships and symmetry, allowing it to better capture local texture patterns and handle variations in pose, illumination, and other challenges. For future work, we will implement the suggested system in videos featuring diminutive facial features in various settings.

References

[1] D. Mamieva, A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Improved face detection method via learning small faces on hard images based on a deep learning approach," *Sensors*, vol. 23, no. 1, p. 502, 2023.

[2] N. Zhang, J. Luo, and W. Gao, "Research on face detection technology based on MTCNN," *Proc. - 2020 Int. Conf. Comput. Network, Electron. Autom. ICCNEA 2020*, pp. 154–158, 2020, doi: 10.1109/ICCNEA50255.2020.00040.

[3] J. Deng, J. Guo, X. An, Z. Zhu, and S. Zafeiriou, "Masked Face Recognition Challenge: The InsightFace Track Report," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2021-Octob, pp. 1437–1444, 2021, doi: 10.1109/ICCVW54120.2021.00165.

[4] Aziz, A. Mirzaliev, S. "Ransomware Threats in Industrial Internet of Things Networks: A Detection Approach," Journal of International Journal of Wireless and Ad Hoc Communication, vol. 8, no. 1, pp. 15-20, 2024. DOI: https://doi.org/10.54216/IJWAC.080102

[5] S. Yang, P. Luo, C.-C. Loy, and X. Tang, "Wider face: A face detection benchmark," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5525–5533.

[6] L. Alzubaidi *et al.*, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. big Data*, vol. 8, pp. 1–74, 2021.

[7] J. Dai, "NIPS-2016-r-fcn-object-detection-via-region-based-fully-convolutional-networks-Paper.pdf," no. Nips, 2016.

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Adv. Neural Inf. Process. Syst.*, vol. 28, 2015.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[10] W. Liu *et al.*, "Ssd: Single shot multibox detector," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, Springer, 2016, pp. 21–37.

[11] Vimala, S. Krishnan, V. Raj, M. Kumar, A. Janakiraman, M. "Object Detection Using Deep Learning," Journal of Journal of Cognitive Human-Computer Interaction, vol. 6, no. 1, pp. 32-38, 2023. DOI: https://doi.org/10.54216/JCHCI.060103

[12] A. S. Joshi, S. S. Joshi, G. Kanahasabai, R. Kapil, and S. Gupta, "Deep Learning Framework to Detect Face Masks from Video Footage," *Proc. - 2020 12th Int. Conf. Comput. Intell. Commun. Networks, CICN 2020*, pp. 435–440, 2020, doi: 10.1109/CICN49253.2020.9242625.

[13] W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, "YOLO-face: a real-time face detector," *Vis. Comput.*, vol. 37, no. 4, pp. 805–813, 2021, doi: 10.1007/s00371-020-01831-7.

[14] N. Kulkarni, D. Mantri, P. Pawar, M. Deshmukh, and N. Prasad, "Occlusion and Spoof Attack Detection using Haar Cascade Classifier and Local Binary Pattern for Human Face Detection for ATM," *AIP Conf. Proc.*, vol. 2494, no. October, 2022, doi: 10.1063/5.0107262.

[15] Y. Liu, R. Liu, S. Wang, D. Yan, B. Peng, and T. Zhang, "Video Face Detection Based on Improved SSD Model and Target Tracking Algorithm," *J. Web Eng.*, vol. 21, no. 2, pp. 545–567, 2021, doi: 10.13052/jwe1540-9589.21218.

[16] E. Solomon, A. Woubie, and E. S. Emiru, "Autoencoder Based Face Verification System," 2023,

DOI: https://doi.org/10.54216/JCIM.140225
Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024

[Online]. Available: http://arxiv.org/abs/2312.14301

[17] L. Wolf, T. Hassner, and I. Maoz, "Face recognition in unconstrained videos with matched background similarity," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 529–534, 2011, doi: 10.1109/CVPR.2011.5995566.

[18] W. K. Mutlag, S. K. Ali, Z. M. Aydam, and B. H. Taher, "Feature Extraction Methods: A Review," *J. Phys. Conf. Ser.*, vol. 1591, no. 1, 2020, doi: 10.1088/1742-6596/1591/1/012028.

[19] B. T. Devi and S. Shitharth, "Multiple Face Detection Using Haar - AdaBoosting, LBP-AdaBoosting and Neural Networks," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1042, no. 1, p. 012017, 2021, doi: 10.1088/1757-899x/1042/1/012017.

[20] A. B. Shetty, Bhoomika, Deeksha, J. Rebeiro, and Ramyashree, "Facial recognition using Haar cascade and LBP classifiers," *Glob. Transitions Proc.*, vol. 2, no. 2, pp. 330–335, 2021, doi: 10.1016/j.gltp.2021.08.044.

[21] B. Yang and S. Chen, "A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image," *Neurocomputing*, vol. 120, pp. 365–379, 2013, doi: 10.1016/j.neucom.2012.10.032.

[22] K. Meena and A. Suruliandi, "Local binary patterns and its variants for face recognition," *Int. Conf. Recent Trends Inf. Technol. ICRTIT 2011*, no. October, pp. 782–786, 2011, doi: 10.1109/ICRTIT.2011.5972286.

[23] A. D. Salman, M. A. Talab, and R. R. Al-Dahhan, "Features extraction for robust face recognition using GLCM and CS-LBP," in *Proceedings of International Conference on Emerging Technologies and Intelligent Systems: ICETIS 2021 Volume 2*, Springer, 2022, pp. 175–191.

[24] Z. Xiong, M. Liu, and Q. Guo, "Finger Vein Recognition Method Based on Center-Symmetric Local Binary Pattern," *Proc. 2019 IEEE Int. Conf. Artif. Intell. Comput. Appl. ICAICA 2019*, pp. 262–266, 2019, doi: 10.1109/ICAICA.2019.8873475.

[25] Abd, B. Obaidi, M. "Enhancing Healthcare Data Classification: Leveraging Machine Learning on ChatGPT-Generated Datasets," Journal of International Journal of Advances in Applied Computational Intelligence, vol. 5, no. 2, pp. 34-45, 2024. DOI: https://doi.org/10.54216/IJAACI.050203

[26] V. H. Phung and E. J. Rhee, "A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets," *Appl. Sci.*, vol. 9, no. 21, p. 4500, 2019.

[27] D. Bhatt *et al.*, "CNN variants for computer vision: History, architecture, application, challenges and future scope," *Electronics*, vol. 10, no. 20, p. 2470, 2021.

DOI: https://doi.org/10.54216/JCIM.140225

Received: January 29, 2024 Revised: April 08, 2024 Accepted: July 12, 2024