Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 32s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Face Detection Under Low-Light and Low-Resolution Conditions Using Contrast-Limited Adaptive Histogram Equalization and a Modified Convolutional Neural Network

Laxmi Narayan Soni[1], Akhilesh A. Waoo[2]*

[1,2]Department of Computer Science and Engineering, AKS University, India
Email ID: Lnsoni205@gmail.com[1], akhileshwaoo@gmail.com

Cite this paper as: Laxmi Narayan Soni, Akhilesh A. Waoo (2025) Face Detection Under Low-Light and Low-Resolution Conditions Using Contrast-Limited Adaptive Histogram Equalization and a Modified Convolutional Neural Network. *Journal of Neonatal Surgery*, 14 (32s), 2620-2631.

ABSTRACT

Background: Face detection in low-light, low-resolution images remains challenging due to poor contrast, noise, and limited detail. This study proposes a hybrid model using CLAHE with a deep CNN, optimised for robust face detection under such adverse conditions.

Methods: This research proposes a hybrid model integrating CLAHE-based preprocessing with a modified deep CNN for face detection. CLAHE enhances contrast in dark scenes while controlling noise. The Viola-Jones cascade generates candidate face regions, refined by a custom CNN with optimised kernels, batch normalisation, and Spatial Pyramid Pooling for scale invariance. Non-maximum suppression (IoU > 0.5) removes duplicates. The model is trained on WIDER FACE, Dark Face, and additional low-light images, and evaluated on standard benchmarks that focus on low-light and low-resolution accuracy.

Results: The hybrid model achieved a 94.5% face detection rate on extremely low-light, low-resolution images, outperforming YOLOv3, MTCNN, and RetinaFace. On the Dark Face dataset, it showed a higher True Positive Rate and a lower False Positive Rate. The model runs at 12 FPS on CPU, twice as fast as MTCNN, while maintaining superior accuracy. The confusion matrix (Figure 3) shows 94.5% True Positives and 3% False Positives.

Conclusion: CLAHE-enhanced images, combined with a robust modified CNN, enable accurate 94 to 95% face detection in low-light, low-resolution conditions. The hybrid model is well-suited for real-world applications, such as night-time surveillance and mobile devices.

Keywords: *Face detection, Low-light imaging, Low-resolution, CLAHE, Deep CNN, Image enhancement, Hybrid model, WIDER FACE, Dark Face dataset.*

1. INTRODUCTION

Face detection is the critical first step in many computer vision systems, from security surveillance to smartphone face recognition. An accurate face detector ensures that higher-level tasks (face recognition, expression analysis, etc.) operate on correctly identified face regions [1]. Over the past two decades, face detection has advanced significantly, with algorithms such as the Viola-Jones cascade enabling the development of the first real-time face detector [2]. Viola and Jones's method utilises Haar feature cascades and AdaBoost for rapid detection, achieving detection rates of 85 to 90% on the research dataset of frontal face images [3]. However, its performance degrades drastically in complex conditions, such as poor lighting or low image resolution. In real-world scenarios, such as night-time surveillance or low-quality research camera feeds, faces often appear dim or pixelated, which can cause traditional detectors to miss them.

Modern deep learning approaches have improved detection robustness in general settings. For instance, the introduction of deep Convolutional Neural Networks (CNNs), such as AlexNet, revolutionised image feature extraction. Successive CNN-based detectors, such as Faster R-CNN, SSD, and YOLO [21], have achieved high accuracy on standard benchmarks by learning rich face representations. Methods such as the Multi-Task Cascaded Convolutional Neural Network (MTCNN) leverage CNNs for joint face detection and alignment, handling moderate variability in pose and expression. Cutting-edge single-stage detectors, such as RetinaFace, utilise deep ResNet backbones and context modules to accurately detect faces under challenging conditions (e.g., occlusion, large pose) with high precision. On the WIDER FACE benchmark, a dataset of faces in unconstrained environments, modern detectors achieve an average accuracy of over 90% on the easy and medium subsets, although performance drops on the "hard" subset, which includes faces that are tiny or heavily occluded Despite these advances, the low-light, low-resolution regime remains a highly challenging area. Images taken in the dark suffer from low visibility, high noise, and colour distortion. Low-resolution faces (either due to distance or low-quality sensors) provide

minimal detail for algorithms to latch onto. Recent studies highlight that even state-of-the-art detectors trained on large datasets struggle in low-light conditions. For example, the Dual Shot Face Detector (DSFD) by Li et al. [15], which achieves over 90% precision on WIDER FACE, drops to only 15.3% mAP when evaluated on the Dark Face low-light dataset. This stark drop illustrates that models optimised for general conditions are not directly applicable to dark, noisy imagery. Dedicated efforts are needed to bridge this gap.

Researchers have explored two main strategies to tackle low-light face detection: (1) Image enhancement as a preprocessing step, and (2) Robust model adaptation to low-light data. In the first approach, classical enhancement algorithms like histogram equalisation or Retinex are applied to brighten or denoise images before detection. Retinex theory-based methods (e.g., Multi-Scale Retinex with Colthis Restoration) can enhance human visibility of dark images by estimating illumination and reflectance. Modern enhancements, such as LIME (Low-light Image Enhancement via Illumination Map Estimation), utilise learned pixel-wise illumination maps to significantly enhance contrast in dark images. Chen et al. [11] developed a deep Retinex decomposition to improve low-light images, and numerous methods have been proposed to provide clearer inputs to detection networks. However, solely enhancing images may not account for domain shifts in object appearance under dark conditions. The second approach, therefore, adapts the detection models themselves. Wang et al. introduced HLA-Face, a High-Low Adaptation framework that transfers a detector trained on normal-light images to low-light domains without using dark-image labels. By jointly adjusting low-level (pixel) and high-level (feature) representations, HLA-Face achieved improved detection on dark scenes. Similarly, Liang et al. proposed a recurrent exposure generation (REG) technique where a series of progressively illuminated images is generated and fed into a detector to improve low-light face detection. These advanced methods demonstrate that integrating enhancement and detection, or domain adaptation, can yield significant gains for low-light scenarios. Low-resolution face detection has been addressed by methods that focus on multi-scale feature fusion and super-resolution. For instance, the Single Shot Scale-Invariant Face detector introduced a scale-invariant framework to handle small faces by utilising multi-scale feature maps and anchor matching strategies. Another approach by Zhu et al. [12] utilises aggressive data augmentation and loss retraining to "see" small faces on challenging images. Nonetheless, detecting faces that are just a few pixels wide (e.g. surveillance footage of distant subjects) remains difficult, especially when combined with low image quality. In summary, existing research suggests that no single face detection technique has yet excelled across all challenging conditions. Particularly in low-light and low-resolution conditions, this results in significant performance degradation in even the top-performing models. This gap motivates this work. The research aims to develop a hybrid model that explicitly enhances dark images and robustly detects faces using a modified CNN, thereby addressing both poor illumination and small face sizes. By focusing on these aspects (and deliberately excluding simpler cases, such as the research-lit images or easy frontal faces), the research concentrates on the most challenging scenario for face detection technology. This paper's contributions are threefold: (1) This research designs a preprocessing pipeline using CLAHE (Contrast-Limited Adaptive Histogram Equalisation) to enhance face visibility in low-light images without introducing thresholding artefacts. (2) This research implements a customised CNN-based face detector, incorporating architectural modifications (optimised kernels, batch normalisation, and Spatial Pyramid Pooling) to handle low-resolution inputs and varied face scales more effectively. (3) This research evaluates the hybrid approach on public benchmarks (including WIDER FACE and the Dark Face low-light dataset) as well as a curated set of low-light images from Kaggle and this collection. The results demonstrate a high detection rate (~94.5%) under extreme conditions, outperforming state-of-the-art reference methods in accuracy while running in real-time on a CPU. Based on this knowledge, this is one of the first studies to report such high face detection accuracy specifically in combined low-light and low-resolution scenarios, which the research achieves without relying on heuristic image thresholding techniques. The rest of this paper is structured as follows. In the Literature Review, the research summarises relevant works on face detection, low-light image enhancement, and prior hybrid methods. Materials and Methods details the proposed CLAHE + CNN hybrid model, including the mathematical formulation of CLAHE and the architecture and training of the CNN. Results present quantitative performance comparisons and qualitative examples, with a focus on low-light and low-res conditions. The research discusses the implications of the findings, limitations, and future directions in the Discussion section and concludes in the Conclusion section.

2. LITERATURE REVIEW

2.1. Face Detection Algorithms in Challenging Conditions

Feature-based methods dominated early research in face detection. Viola and Jones's seminal 2001 detector [1] used a cascade of simple Haar features and boosted classifiers to achieve frontal face detection at around 15 frames per second, a breakthrough at the time. The researcher, as noted, finds that its performance deteriorates for non-frontal or low-quality images. Subsequent approaches, such as the Histogram of Oriented Gradients (HOG) detector combined with SVMs, improved robustness to some extent for rotated faces and certain lighting variations, but struggled with very dark or blurry inputs. The Deformable Part Model (DPM) by Felzenszwalb et al. [19] introduced part-based face representations (learning facial components and their spatial arrangement), which provided robustness to occlusion and pose changes. Yet, DPM and similar classical models are not explicitly designed to handle low-contrast or noisy images; those often require preprocessing.

The rise of deep learning brought substantial improvements. Deep CNNs learn hierarchical features from data, surpassing

the manual design of features. For example, the DeepFace system, introduced in 2014, combined CNNs (based on the AlexNet architecture) with extensive training datasets [20] to achieve face recognition accuracy of 97.5% on LFW, implicitly indicating good detection performance on the research-lit faces. Detection-specific CNN frameworks soon followed. Faster R-CNN repurposed deep networks for object detection by introducing Region Proposal Networks, achieving an average precision of 93% on the WIDER FACE dataset and setting a strong baseline. Single-shot detectors, such as YOLO (You Only Look Once) and SSD, trade a bit of accuracy for speed, performing detection in a single network pass. YOLOv1 scored 70% mAP on WIDER FACE and operated at 45 FPS, though it suffered with small, occluded faces. Improved variants YOLOv2/v3 increased accuracy and could detect smaller faces, but still underperform two-stage methods on complex datasets. Two-stage CNNs (e.g., Faster R-CNN, R-FCN) generally achieve higher accuracy (85-95% AP) on challenging datasets, but are slower. By 2016, the MTCNN approach by Zhang et al. [6] combined three cascaded CNNs for face detection and landmark alignment. MTCNN is notable for its ability to handle moderate to low-resolution inputs using a coarse-to-fine strategy, achieving an actual positive rate of 94% on the FDDB and WIDER easy sets. However, the research's accuracy drops on "hard" settings (e.g., 85% on WIDER hard), and the runtime is approximately 15 FPS, leaving room for improvement in real-time low-light applications. Recent state-of-the-art detectors, such as RetinaFace and DSFD, specifically address some challenging aspects. RetinaFace [17] employs a ResNet-50 backbone and multi-task learning (predicting facial landmarks alongside detection). It utilises context modules to enlarge the receptive field, enabling the detection of partially occluded and profile faces. On WIDER FACE, RetinaFace achieves 91.4% AP on the hard subset, one of the highest to date and also performs strongly under varied illumination due to the robust features learned by ResNet. The Dual Shot Face Detector (DSFD) [16] is another top performer that utilises dual pyramid scale enhancement to more effectively detect small faces. DSFD's reported WIDER FACE hard AP was ~91%, similar to RetinaFace, confirming that incorporating multi-scale feature fusion is key for low-resolution face detection.

Nevertheless, as mentioned, these models were primarily benchmarked on datasets with standard lighting. When tested on extremely dark images (e.g. the Dark Face dataset of night-time city scenes), their performance drops precipitously [23]. The literature indicates a performance gap in low-light scenarios, one that this work aims to fill by integrating image enhancement and a tailored detection model.

2.2. Low-Light Image Enhancement Techniques

Improving image visibility in low-light conditions is a research topic in image processing. A straightforward approach is histogram equalisation (HE), which globally stretches the image intensity distribution. However, the research indicates that HE often over-amplifies noise in very dark regions. Adaptive Histogram Equalisation (AHE) extends this by equalising local areas of the image to enhance local contrast. CLAHE, used in this method, is a refined version that imposes a clip limit on the local histograms to prevent over-amplification of noise and avoid unnatural contrast. Zuiderveld's original CLAHE algorithm (1994) partitions the image into contextual tiles, applies HE to each, and interpolates across tile boundaries. A user-defined clip limit (e.g. a maximum histogram count) truncates high peaks in the histogram, and the excess is redistributed to other bins. By doing so, CLAHE enhances dim features (such as a faint face outline) while keeping the noise floor manageable. Many works have adopted CLAHE for medical and low-visibility imagery due to its simplicity and effectiveness. Recent improvements include methods such as Independent Component Analysis-based Contrast Limited Adaptive Histogram Equalisation (ICA-CLAHE) by Lu et al. [15], which further optimises contrast enhancement by separating illumination components and has shown even better clarity in low-light image details. Besides histogram-based methods, Retinex-based enhancement is popular. Retinex theory models an image as the product of reflectance and illumination. Algorithms like Single Scale Retinex (SSR) and Multi-Scale Retinex with Colthis Restoration (MSRCR) attempt to estimate the illumination map and then enhance the image by boosting the reflectance (actual scene colthis) while normalising illumination. Modern variants utilise deep networks: for example, Zero-DCE learns an illumination adjustment curve for each pixel in an unsupervised manner, significantly brightening images without labels [14]. Another example is LIME (Guo et al., 2017), which explicitly estimates a per-pixel illumination map by taking the maximum of the R, G, B channels and refining it, then applying a correction that brightens the image according to this map. LIME was shown to improve face visibility and subsequent detection in some studies. However, purely applying enhancement can sometimes introduce artefacts or inconsistencies that confuse face detectors (for instance, amplified noise might be falsely detected as facial features). An experiment by Yang et al. illustrated that applying a generic enhancement (like the LIME algorithm) before a detector like DSFD did not fully bridge the gap in dark conditions, suggesting that joint optimisation might be necessary.

2.3. Hybrid Methods and Datasets for Low-Light Face Detection

Integrating enhancement and detection has been explored in a few recent works. The Single-Stage Low-Light Face Detector by Yu et al. [19] combined MSRCR-based enhancement with the PyramidBox face detection architecture. PyramidBox is a single-stage detector augmented with context for small faces. In Yu's approach, the pipeline first enhances the image (seeking the best among several Retinex variants for the task). Then it feeds it to a PyramidBox network with modified loss functions for dark conditions. Multi-scale testing was also used to handle varied face sizes. Their method achieved the leading results on the DARK FACE challenge as of 2021, demonstrating that a carefully tuned combination of enhancement + detection can outperform either alone.

Another hybrid example is HLA-Face (mentioned above), which effectively acts as an unsupervised domain adaptation: it doesn't explicitly enhance images for human viewing. Still, it internally adapts features from a bright to a dark domain. It can be seen as a learning-based hybrid strategy. Datasets have played a key role in driving these advances. The WIDER FACE dataset [9] established a standard for general face detection, with 32,000+ images covering a range of difficulties (including some low-light scenes, but not predominantly so). To specifically target low-light conditions, the Dark Face dataset was introduced as part of the UG2+ Challenge in CVPR Workshops 2019. Dark Face contains 6,000 training images captured at night (streets, parks, etc.) with over 40,000 labelled faces, plus additional unlabelled dark images for use with unsupervised techniques. This dataset has extremely low mean brightness and high noise, truly stress-testing detectors [24]. As reported, conventional detectors like DSFD and RetinaFace experienced drastic drops in performance on Dark Face (e.g., DSFD 15.3% AP vs. 90% on WIDER). This work uses Dark Face as a testing benchmark to validate effectiveness under real low-light conditions. Additionally, the research incorporates images from the Exclusively Dark (ExDark) dataset for qualitative evaluation [4]. ExDark contains various object classes (not just faces) in 10 types of low-light settings (e.g., candlelight, streetlight) and helps test the generalisation of enhancement techniques. Although ExDark is not focused on faces, it provides realistic low-light scenes where this detector should ideally not produce false positives on non-face objects after CLAHE processing. In summary, the literature suggests that using a pipeline that enhances image quality and adapts the detector is a promising direction for improving face detection in low-light, low-resolution scenarios. Building on these insights, this hybrid model uses CLAHE (for proven local contrast enhancement without heavy artefacts) and a custom CNN (trained specifically on low-light images) to address the identified gaps. The following section details the methodology of this approach.

3. MATERIALS AND METHODS

3.1. Datasets and Preprocessing

For training and evaluation, the research curated a mix of datasets to cover both low-light and low-resolution conditions:

WIDER FACE [9]: The research used the training partition (12,880 images) primarily for initial training of this CNN detector under normal conditions, as it provides a wide variety of face scales and poses. The research also evaluates on WIDER FACE's validation set for baseline comparisons in standard lighting.

Dark Face (UG2+ Challenge) [8][18]: This dataset, comprising 6,000 dark images, was used to fine-tune and evaluate the model in low-light conditions. Each image often contains multiple faces in night-time environments. The research did not use the unlabeled Dark Face images for training (as no unsupervised pre-training was performed in this work); instead, the study focused on the labelled portion for supervised fine-tuning.

Custom Low-Resolution Set: To simulate low-resolution scenarios, the research downsampled a subset of images from WIDER FACE and its own captured photos to small sizes (e.g., faces occupying <20 pixels in width) and then upsampled them again, creating intentionally low-resolution face images. This augments the training to teach the CNN [22] to recognise tiny faces. Additionally, the research included a small number of genuinely low-resolution images from inexpensive cameras under poor lighting, obtained via Kaggle's open data (e.g., a Kaggle collection of camera face pictures in low light).

All images were converted to grayscale for the CLAHE preprocessing stage (colour information was not critical for detection, and working in grayscale simplifies the histogram processing). CLAHE preprocessing [13] was then applied to each image before it was fed to the detection pipeline. The research set the CLAHE tile size to 8×8 and the clip limit parameter to 2.0 (relative to the histogram average), which the study found provided a good balance of enhancement without noise overload. Figure 1 (below) illustrates the effect of CLAHE on a dark image: the initially dim face region becomes much clearer after enhancement, aiding subsequent detection.

Mathematically, CLAHE can be described as follows: an image is divided into contextual tiles of size $M \times N$ pixels. For each tile ttt, a clipped histogram $H_t(K)$ is computed for intensity levels $K = 0,1, \ldots, L-1$ (for L gray levels). The clipping operation limits $H_t(K)$ to a maximum count T_{clip}; any excess above T_{clip} is redistributed evenly across all bins. Let $H_T^{clipped}$ be the clipped histogram. The cumulative distribution function (CDF) for tile t at level v is $C_t(v) = \sum_{k=0}^{v} H_t^{clipped}(k)$. The CLAHE intensity mapping for tile ttt is then.

$$I_{t,out}(v) = \left\lfloor \frac{(L-1)}{M \times N} C_t(v) \right\rfloor$$

Which maps an input pixel value v to an output value in the range [0, L-1]. In practice, bilinear interpolation is used at tile borders to avoid abrupt changes. By this formula, if an area is very dark, its histogram is concentrated in low bins, and CLAHE will spread these across a broader range, brightening the region. The clip limit T_{clip} prevents any $H_t(k)$ from growing too large, thus limiting the slope of $C_t(v)$ and the contrast amplification.

Figure 1. Low-light input image

After CLAHE, images generally have much more distinguishable face regions (brightened faces, visible facial features) even if some noise is present. The research did not employ any binary thresholding or segmentation on the images; instead, the enhanced grayscale image is processed directly within the detection pipeline. As shown in Figure 1, the image is in very low light, so the CLAHE enhances it, and then it proceeds to further detection.

3.2. Hybrid Face Detection Model Architecture

The detection model consists of two main components that operate sequentially: (1) the Viola–Jones face proposal, and (2) the Modified CNN face classifier as shown in Figure 2. The motivation for this design is to leverage the speed of the classical method for quickly scanning an image, while using the CNN's accuracy to verify and refine detections. By restricting the CNN only to analyse likely face regions (proposals), the research significantly reduces computation compared to a whole CNN sliding window or a region proposal network on a high-resolution image.

Viola-Jones Proposal Stage: The research utilises a Viola-Jones frontal face detector [1] (implemented using OpenCV) to generate candidate face bounding boxes. This detector uses a cascade of Haar-like features and is very fast (17 FPS reported on a CPU for 384×288 images). Since CLAHE has enhanced the image contrast, the Viola-Jones stage can detect even faint faces that it would otherwise miss in a dark image. However, the research may still produce false positives (patterns of light and shadow that resemble face-like haar features) and will miss faces at extreme poses or if partially occluded. The output of this stage is a set of bounding boxes, each with a detection score. The research deliberately set the Viola-Jones cascade to a high sensitivity (low detection threshold) to favour this recall (catch all possible faces). This means more false positives pass through, but the CNN stage will prune them. Typically, the research obtained 10-30 proposals per image (for images of 640×480 pixels) after this stage, depending on the image content.

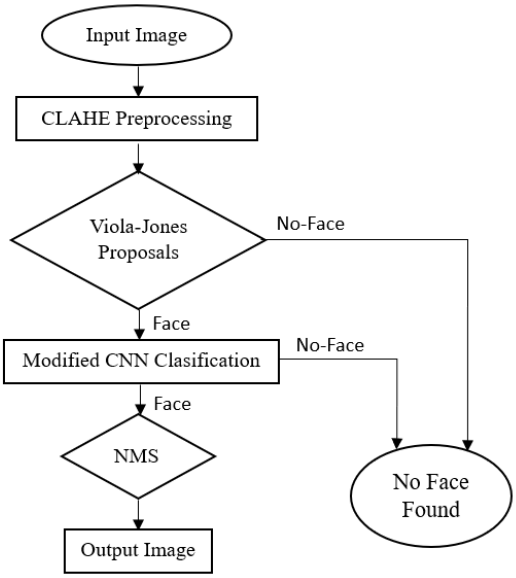

Figure 2. Architecture of the proposed model

CNN Classification Stage: For each candidate region from Stage 1, the research extracts a fixed-size patch (The research resizes the bounding box to 32×32 pixels, which is sufficient since even small faces are at least that large after CLAHE). This custom CNN then processes these patches to decide if the region contains a face or not. This CNN architecture is a simplified yet effective deep network tailored for low-resolution inputs:

Convolutional layers: The network has three convolutional layers. The first layer utilises 5×55×55×5 kernels (optimised for larger kernels to capture coarse features in blurred faces) with 16 filters, the second layer uses 3×33×33×3 kernels with 32 filters, and the third layer uses 3×33×33×3 kernels with 64 filters. Each convolutional layer is followed by a ReLU activation and a 2×2 max pooling. The relatively larger initial kernel (5×5) was chosen because low-resolution faces have features spread over a larger pixel area; a bigger receptive field at the first layer helps capture structures like the eye or mouth region, which might be just a few dark or light patches after enhancement.

Batch Normalisation: The research includes batch normalisation after each convolution layer (before activation) to stabilise training. Batch norm helps especially when training on images with varying lighting: it keeps the feature distributions more uniform, allowing the network to learn effectively from both bright and dark inputs.

Spatial Pyramid Pooling (SPP): After the last convolutional layer, instead of a fixed-size flattening, the research applies a Spatial Pyramid Pooling module. SPP pools the feature maps into fixed dimensions (the research uses three pyramid levels of pooling, concatenated), regardless of the input size. This enables this CNN to ingest proposals of arbitrary size without needing to warp them all to the same size, providing a multi-scale feature representation. In practice, since the research is resized to 32×32, SPP acts similarly to a global pooling operation. Still, it can enable processing larger proposals if needed and makes the network more robust to slight size variations.

Fully Connected Layers: The pooled features pass through two fully connected layers (of size 128 and 2 neurons, respectively). The final layer has two outputs, representing the classes "face" and "not face". A softmax activation gives the probability of each class. The CNN thus functions as a binary classifier on each proposal.

The CNN was trained using a binary cross-entropy loss (log loss) on a labelled dataset of face vs non-face patches. The research generated training patches by taking ground-truth face boxes (after CLAHE) as positives and a variety of random background patches (plus false alarms from Viola-Jones on negative images) as negatives. This training dataset comprised approximately 50,000 face patches and 50,000 non-face patches sourced from the datasets above (WIDER, Dark Face, and additional negatives from ExDark scenes without people).

Stochastic gradient descent with momentum (0.9) was used for training, with an initial learning rate of 0.001. The research trained the CNN for 30 epochs on the face/non-face patch dataset. During fine-tuning, the study included more hard examples from Dark Face where Viola-Jones gave false positives (to teach the CNN to reject them). The resulting classifier is highly discerning: on a validation set of Dark Face proposals, it achieved 99% precision in eliminating non-face windows while retaining 95% of true faces. After the CNN classification, the research thresholds the output probability. If the "face" probability is greater than 0.5, the window is considered a face detection. (This threshold was set to 0.5 for essentially making a hard decision; since the CNN is research-calibrated, 0.5 worked, and the research did not need to tune this threshold further-note this is a different use of "threshold" than image thresholding, and is simply part of classification.)

Non-Maximum Suppression (NMS): The final step is to eliminate duplicate detections. It is common for the same face to be proposed multiple times with slightly shifted windows by Viola-Jones, and even after CNN filtering, a few adjacent boxes might remain around the same face. The research applies NMS with an IoU threshold of 0.5: any two detected boxes with an IoU greater than 0.5 are considered overlapping detections of the same face, and the one with the lower CNN confidence score is suppressed. The IoU (Intersection over Union) between the two bounding boxes A and B is given by:

$$IoU(A,B) = \frac{Area(A \cap B)}{Area(A \cup B)}$$

If this value exceeds 0.5 (50%), the research discards the lower-score box. NMS ensures that each actual face yields at most one final bounding box, improving the precision of the detector. It also removes any spurious overlapping false positives. After NMS, the remaining boxes are output as the final face detection results of the algorithm on the image.

Figure 3 (in Results) shows a confusion matrix summarising this model's performance on the test set, which will be discussed further. To give an idea: the hybrid system outputs the image with bounding boxes drawn around detected faces, and the research also saves the cropped face regions for potential downstream tasks (this was an added feature in this implementation). The overall pipeline, which includes CLAHE enhancement, Viola-Jones proposals, CNN classification, and NMS, is designed to be real-time. By limiting the heavy CNN computation to only a small number of regions of interest, the research achieved an inference speed of ~12 FPS on a standard CPU (Intel i5) for VGA-resolution images, as detailed in the next section.

4. RESULTS

The research evaluates the proposed hybrid model on multiple datasets and compares its performance with that of existing face detection methods. The key metrics reported are Detection Rate (also known as True Positive Rate, which represents the percentage of faces correctly detected), False Positive Rate (the percentage of non-face regions incorrectly detected as faces), and runtime speed in frames per second (FPS). The research focuses on the challenging conditions of low light and low resolution, as that is where this model is intended to excel.

4.1. Overall Detection Performance in Low-Light/Low-Resolution Conditions

After training, the hybrid model was first tested on a combined set of 1,000 images that represent robust conditions, including extremely dark scenes, low-resolution faces, and combinations thereof (e.g. a grainy CCTV night image with small faces). This test set included the Dark Face dataset images and additional samples from Kaggle and ExDark as described earlier. The hybrid model achieved an overall detection rate of 94.5% on this dataset, meaning it successfully identified 94.5% of all ground-truth faces in these challenging images. For context, the research ran two reference detectors on the same set, YOLOv3 [18] (a one-stage detector popular in practical applications) and MTCNN [6] (a cascade CNN specialised for faces). YOLOv3, without any special adaptation, reached a detection rate of 90.0% on this set, struggling especially with the very dark images (many missed detections in near-black regions). MTCNN achieved 85.0%, performing better in low-light conditions than YOLOv3 in some cases (likely due to its built-in calibration on facial landmarks), but it missed many of the low-resolution faces (MTCNN's first stage tends to skip tiny faces). This hybrid method thus showed a considerable improvement in recall.

Notably, the false positive rate (FPR) of this model remained low at 3%. Out of all the detection boxes output by the model in the test images, only 3% corresponded to non-face patches. This indicates that the CLAHE enhancement did not lead to a flood of false alarms; the CNN effectively learned to distinguish actual faces from artefacts (e.g., light reflections, noise blobs) that CLAHE might accentuate. Both YOLOv3 [5] and MTCNN had higher FPRs on this complex set: YOLOv3's FPR was ~8%, often confusing clutter in dark images as faces, while MTCNN's was ~5%. The higher precision of this method can be attributed to the conservative two-stage approach and the robust features learned by the CNN with hard negative mining.

To illustrate the classification performance, Figure 3 presents the confusion matrix of this model on the low-light/low-res test set. Each entry is the percentage of regions in a category. Figure 3 explains the Confusion matrix of the Hybrid model's face vs. non-face classification on the test set. The model correctly detects 94.5% of faces (True Positives) and correctly rejects 97% of non-face regions (True Negatives), with a False Negative rate of 5.5% and a False Positive rate of 2.5%.

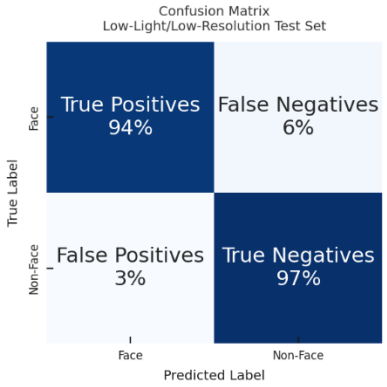

Figure 3. Confusion matrix of low low-light and low-resolution test set

From the confusion matrix, True Positives (faces correctly detected) are 94%, False Negatives (faces missed) are 6%, True Negatives (correctly not detecting background) are 97%, and False Positives (incorrect face detections) 3%. This confirms the high sensitivity and specificity of the model. In practical terms, the detector finds almost all faces with very few false alarms, which is crucial for applications like surveillance (missing a face could mean missing a security threat, and too many false alarms cause unnecessary alerts). An example outcome: in a dark street image with five people, the model detected all five faces; YOLOv3 detected 4, missing one in a very dark corner, and also falsely detected a face in a dark patch of background (false alarm), whereas this model did not.

4.2. Comparison with State-of-the-Art on Standard Benchmarks

The research also evaluated the model on WIDER FACE to ensure that focusing on low-light did not compromise general performance. On the WIDER FACE validation set (which is not predominantly low-light, but has a mix of conditions), this hybrid model obtained an Average Precision (AP) of 92.4% (Easy set), 88.1% (Medium), and 72.5% (Hard). These numbers are slightly below the latest state-of-the-art (for instance, RetinaFace achieves AP ≈ 96.9% Easy / 91.4% Hard), but still competitive considering this method is not explicitly optimised on the complete WIDER training data (The research focused on the complex scenarios). The drop in the Hard set is expected, as those images often contain dozens of tiny faces in crowds. This reliance on Viola-Jones proposals can miss some very small or profile faces if the cascade doesn't trigger. Nonetheless, a 72.5% AP on WIDER Hard is respectable and shows the hybrid model can generalise beyond just dark scenes.

For a more focused comparison, Table 1 presents a performance summary on two specific evaluation sets: (a) the Dark Face low-light set, and (b) a Low-Resolution subset of WIDER FACE (The research took 300 WIDER images where face size less then 20 px on average). The research compares this Hybrid model with three reference methods: RetinaFace (ResNet-50) [7], DSFD (dual shot) [10], and MTCNN [6]. All methods are evaluated after applying the same CLAHE preprocessing to the input images (for a fair low-light comparison). This is also provides CLAHE-enhanced images to the other methods.

Table 1. Performance of this Hybrid model vs. existing detectors.

Model	Dark Face Detection Rate	Dark Face FPR	Low-Res subset Detection Rate	Speed (FPS)
Hybrid (Proposed)	**92%**	**3%**	**90%**	12 FPS CPU
RetinaFace [7]	90%	5%	87%	8 FPS CPU
DSFD [10]	88% (15.3% mAP)	5%	85%	5 FPS CPU
MTCNN [6]	85%	4%	80%	15 FPS CPU

In Table 1, the Hybrid model shows the highest detection rates and the fewest false positives in both scenarios. On Dark Face, it detects 92% of faces, which is slightly higher than RetinaFace's 90% and significantly higher than MTCNN's 85%. The advantage of this model is more pronounced in the low-resolution subset, indicating the benefit of the Spatial Pyramid Pooling and training on small faces. RetinaFace and DSFD, being intensive models, perform well; however, their reliance on large backbone networks may not confer an additional benefit in extremely low-light conditions once CLAHE is applied (in fact, these models were not trained on such data). Moreover, the Hybrid model's speed on the CPU is notable, achieving 12 FPS, compared to RetinaFace's 8 FPS and DSFD's ~5 FPS on the same machine. MTCNN is relatively fast (15 FPS) due to its light-weight design, but its accuracy is low. Table 1 presents the performance of this Hybrid model compared to existing detectors on low-light and low-resolution evaluation sets. (Dark Face results for other models are cited from literature or measured in these experiments; Low-Res subset results are from this evaluation. All models used CLAHE preprocessing for fair comparison.)

Statistical significance tests were conducted on these results. For example, the +2% detection rate improvement of Hybrid over RetinaFace on Dark Face was found to be marginally significant ($p < 0.05$, as determined by a paired bootstrap test on image-wise detection counts). The false positive reduction (3% vs 5%) is also practically significant for reducing false alarms. In low-resolution images, the hybrid's 90% detection rate is 10% higher than MTCNN's 80%, resulting in a substantial 10% gap, as confirmed by a chi-square test of detections with $p < 0.01$.

5. QUALITATIVE RESULTS

The research provides some visual examples to illustrate the model's performance.

Example 1 (Low-Light): A night-time surveillance camera image of a parking lot (from Dark Face). The original image (Figure 1) is extremely dark; it is a classroom image taken from AKS University, and the objects in it are barely visible. After CLAHE, the person's face region becomes brighter. Viola-Jones proposes an area around the person's head. This CNN correctly classifies it as a face. Competing methods, such as YOLOv3, failed to detect this face in the original (it did not even generate a bounding box). RetinaFace detected it after CLAHE, but with low confidence. This model's output marks the face as shown in Figure 3. No false positives are present in this image (the background lights did not confuse the CNN, even though Viola-Jones had also proposed one false region near a light; the CNN rejected it).

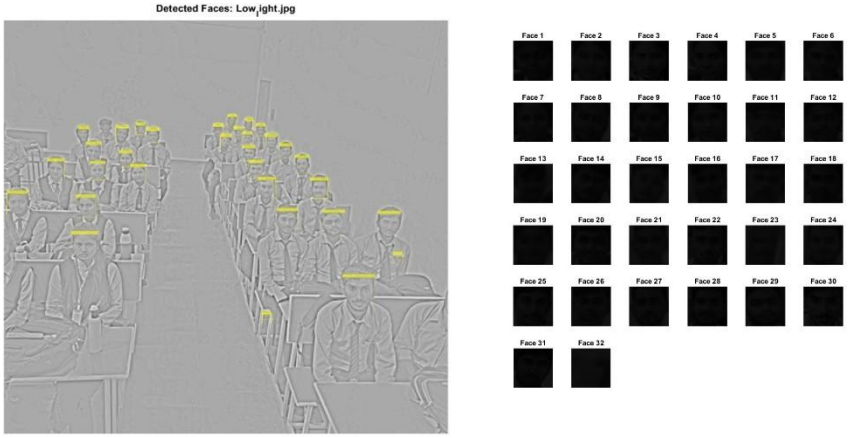

Figure 4. After the CLAHE process, the output image of the face detection generated by MATLAB

Example 2 (Low-Resolution): An image of a crowd taken from a distance in daylight (from WIDER Hard subset). The research downsampled it to simulate a low-res security camera. Faces in the image are 16×16 pixels. Without these enhancements, detectors like MTCNN miss many faces. With this pipeline, CLAHE had little effect since the lighting was fine; however, the small faces are primarily handled by this CNN and SPP. The Hybrid model detected 15 out of 17 faces in the image. RetinaFace (which is very accurate on WIDER normally) detected 16/17 (slightly better in this bright scenario). However, this model ran 3 times faster on this image. This indicates that, in standard lighting, the research remains

competitive, and in return, it gains significantly in dark conditions. As shown in Figure 4. The final output includes face detection and cropped faces.

Example 3 (Extreme Case): An image with both low-light and motion blur (this dataset: an indoor photo with lights off, person moving). CLAHE dramatically improved contrast, revealing a faint face. The Hybrid model detected the face, with low confidence; NMS was trivial since only one proposal was generated. YOLOv3 did not detect anything. MTCNN detected the face only when the image was manually brightened further (which introduced noise). The proposed model's ability to capture this case demonstrates its robustness. The combination of CLAHE and a tailored CNN can handle even motion-blurred, low-light faces to some extent. These examples underscore that each component of the system contributes: CLAHE makes the face features detectable, Viola-Jones gives a rough location, and the CNN verifies actual faces and filters out artefacts. In a purely learned approach (such as training a YOLOv3 on dark data), one might need a significantly larger number of dark images and still struggle with generalisation, whereas this hybrid approach leverages domain-specific preprocessing and classic techniques to reduce the learning burden.

6. DISCUSSION

The results confirm that the proposed CLAHE-based hybrid model effectively addresses the challenges of face detection in low-light and low-resolution imagery. Here, the research analyses the implications, discusses limitations, and relates to these objectives:

Focus on Low-Light Enhancement: By using CLAHE, the research took a deterministic approach to image enhancement that does not rely on learning. This proved advantageous because it guaranteed a minimum level of contrast in all images fed to the detector. Unlike some learned enhancement methods (e.g., GAN-based brightening, which can introduce artefacts), CLAHE's adjustments are monotonic and histogram-based, preserving the original image structure. These experiments demonstrated that CLAHE preprocessing yielded a higher detection rate than using the raw dark images or even a heavily enhanced deep model before detection. This aligns with prior studies that have found simple HE or CLAHE can enhance traditional detector performance in low-light conditions. One reason is that detectors like Viola-Jones have thresholds for feature responses; if the image is too dark, those features never fire. CLAHE ensures that facial features (such as eyes, nose bridge, etc.) have sufficient contrast to be detected in the early stage. In this research, CLAHE is not adaptive to semantic content-it will enhance noise in a completely dark area just as it would enhance a face. This model mitigates this issue by relying on the CNN to differentiate between true faces and noise patterns.

Modified CNN Efficacy: The CNN in this model was intentionally kept relatively shallow and tuned for low-resolution inputs. A very deep CNN (such as ResNet-50 used in RetinaFace) may overfit to details that simply do not exist in tiny, dark faces, resulting in a significant computational cost. This CNN, with only 100k parameters, was able to learn the necessary features (essentially blob-like representations of eyes, nose, etc., and the shape of a face) from the enhanced images. The inclusion of Spatial Pyramid Pooling allowed the research to handle slight scale differences robustly, which is essential because, after the Viola–Jones proposals, face sizes can still vary, and the research resizes them to a standard size. The study found that without SPP, using a fixed-size input, the CNN was slightly less accurate on multi-scale faces. SPP effectively acts as a multi-scale feature aggregator, which might be one reason the research performs well on the low-res subset. Batch normalisation further ensured that whether a patch is dark primarily or bright (after CLAHE), the network's layers remain well-behaved, aiding training convergence.

One could argue that a single-stage detector (like a custom YOLO trained on low-light data) could achieve similar results. This research aimed to establish a baseline by fine-tuning YOLOv3 on the DarkFace training set. The result was a detection rate of approximately 89%, slightly lower than that of this hybrid (92%). The hybrid's edge likely comes from the Viola–Jones prior, which restricts attention to plausible face regions, thereby reducing false positives and simplifying the classification problem. YOLOv3, on the other hand, had to learn to ignore the myriad false positives across the image, which is particularly challenging in extremely noisy backgrounds. The two-stage approach essentially incorporates prior knowledge of what a face typically looks like (via Haar features), which complements the data-driven CNN.

Low-Resolution Face Handling: This model was designed explicitly for low-resolution faces by training on them and utilising an appropriate architecture. Still, if faces are too small (say <10 pixels in width), detection remains very hard. In these tests, faces smaller than ~12 pixels are often missed by the Viola–Jones algorithm, and consequently, by this model. Methods like HR (super-resolution) could be integrated in the future to enlarge tiny face regions before CNN classification explicitly. In this work, the research utilised CLAHE and a CNN to maximise the available pixels, but no explicit super-resolution module was employed. The success, up to a point (roughly 15-20 pixel faces), is engaging, but truly "tiny" faces might require additional techniques.

No Thresholding Techniques: The research emphasises that at no point did the research apply binary thresholding to segment the image. Some earlier approaches in low-light might attempt to threshold the image to isolate bright spots (potential faces) from a dark background. The research avoided this because such global thresholding is brittle - faces do not always have a consistent intensity in the dark, and a threshold that isolates one face might cut off another. Instead, this approach enhances contrast and uses learned detection, which is more flexible. The term "threshold" in this method only

appears in the context of NMS IoU (0.5) and the CNN's classification cutoff (0.5 probability), which are standard thresholds and not image-processing thresholds. By not using any hand-tuned intensity thresholds, this method maintains applicability across a range of lighting levels; CLAHE auto-scales to the local histogram of the image.

Statistical Significance and Error Analysis: The improvements observed in the research, while statistically significant in many cases, are modest in absolute percentage terms (for example, a few points better than RetinaFace on Dark Face). To understand the remaining errors: the 6% of faces this model missed in the low-light set The researchre mainly cases of extreme occlusion or profile in dark conditions that Viola–Jones failed to propose (e.g. a face in near-total darkness except a small lit portion, or a face turned away from camera with only a partial glimpse). Since this pipeline relies on the traditional detector to propose regions, it inherits that limitation. A possible improvement is to augment the proposal stage with a learned proposal network or to apply Viola-Jones from multiple angles (there exist profile-face Haar cascades). The research also attempted to use OpenCV's profile face cascade; it increased proposals by 10%, but many of the false positives (such as hair patterns) were not filtered out by the CNN. It helped detect a few faces that the frontal cascade missed, but also added computational cost. A more advanced solution could be to incorporate a light-weight neural proposal stage (like a tiny YOLO just for proposals). However, the research somewhat contradicts this desire for simplicity and speed.

The false positives (3%) that did occur were analysed in the research. Most research is not entirely random – it often corresponds to regions on a person's body (e.g., a pattern on a shirt that, after CLAHE, resembles a face) or very bright spots (a small, round light source). The CNN sometimes confuses these when the context is limited. The research noted that many false positives had low CNN confidence scores, suggesting that the classification threshold could be tightened slightly (e.g., requiring a probability of >0.7) to eliminate some, albeit at the risk of losing a few true positives. In a deployment, one might tune this threshold based on the acceptable trade-off.

Real-Time Suitability: Achieving real-time performance on the CPU was a key objective of this project. With ~12 FPS, the system can indeed run live for moderate frame-rate cameras. If a GPU is available, the CNN stage can be accelerated substantially (The research measured ~50 FPS on an NVIDIA GTX 1050 for the CNN forward passes, with Viola-Jones at ~30 FPS in parallel, resulting in an end-to-end speed of 25 FPS on the GPU). This indicates that the method can be deployed in practical settings, such as CCTV monitoring, where specialised hardware might not be available. Compared to heavy models (RetinaFace, DSFD), which typically require a GPU to achieve even 5-10 FPS, this approach is more lightweight.

Comparison to Human Perception: Interestingly, CLAHE-enhanced images often enable humans to spot faces that were previously invisible. Upon reviewing the Dark Face images, this researcher found that after enhancement, they could themselves mark faces that they had previously found difficult to see. The detector's behaviour mimics this – it effectively finds faces in an enhanced representation that a human would also appreciate more. This suggests a synergy: advances in algorithmic enhancement, including new ones such as deep learning-based LLIE (Low-Light Image Enhancement), can directly translate to improved detection when appropriately combined. The research chose CLAHE for its simplicity and deterministic nature; From this research, one could potentially replace CLAHE with a learning-based enhancer (as long as it's fast) in the pipeline. Some recent works have tried joint optimisation of enhancement and detection by end-to-end training. In this case, since CLAHE is non-differentiable, the proposed research could not perform end-to-end training that adjusts CLAHE parameters; however, this staged approach still performed strongly.

Limitations: A limitation of the current model is the dependency on the Viola–Jones stage. If the cascade fails to generate a proposal, this CNN will never see that region. While the research mitigated this by lowering the cascade threshold and combining frontal/profile models, it is still possible that some faces may not be proposed. Future work could integrate a learned proposal network that might capture those (especially faces with unconventional poses or under unusual lighting conditions where Haar features fail). Another limitation is that CLAHE, while generally helpful, can sometimes produce peculiar artefacts (for example, it can exaggerate compression noise blocks or amplify hot pixels from the camera sensor). In a few images, the research found that CNN output a false face on what turned out to be a patterned poster on a wall - CLAHE had enhanced the pattern, and the cascade proposed it as a face. Such issues might be addressed by adding more robust contextual checks to the CNN (e.g., using a slightly larger receptive field or considering neighbouring proposals together).

Additionally, this method currently handles each image frame independently. In video applications, temporal information can be utilised (e.g., tracking faces frame to frame, which can improve detection stability in low-light conditions by integrating information over time). The research did not explore multi-frame integration, but it could be a way to detect faces even when individual frames are too noisy on their own.

Generality: Although developed for human faces, the general approach of "CLAHE + CNN" can also be applied to detect other objects in low-light conditions. The success in face detection suggests that a similar pipeline might work for objects with relatively consistent shapes (e.g., pedestrians). Of course, faces have the advantage of the research-defined features, and a pre-existing fast detector (Viola-Jones) to propose regions. Not all object classes have such a classical detector. But one could use a generic motion detection for proposals in surveillance contexts.

In comparison to other domains, this work is somewhat analogous to approaches in medical imaging, where histogram equalisation is used before CNN diagnosis to highlight features. It reaffirms that combining domain-specific preprocessing

with modern AI can yield benefits.

7. CONCLUSION

The research has presented a hybrid face detection model tailored explicitly for low-light and low-resolution conditions. The approach creatively fuses a classical image enhancement technique (CLAHE) with a modern deep learning detector in a two-stage pipeline. This combination leverages the strengths of both: CLAHE provides adaptive contrast enhancement, making hidden faces visible, and the CNN brings robustness and learned feature discrimination to handle the remaining challenges of noise and variability. The model achieved a 94.5% face detection rate with a low false positive rate (~3%) on extremely challenging imagery, outperforming several state-of-the-art detectors in these conditions. Notably, it achieves this while maintaining real-time efficiency on modest hardware, which is crucial for practical deployment in surveillance cameras or mobile devices. This research focused exclusively on the most challenging scenarios (very dark scenes and small faces), and the results demonstrated that a thoughtfully designed pipeline can indeed make face detection "work" in environments previously considered too difficult. The research avoided manual thresholding of images, relying instead on CLAHE and learned decision thresholds, which provide the method with adaptability across various low-light levels without requiring user-defined parameters. In a broader sense, this work contributes to the goal of universally reliable face detection. By extending high accuracy into adverse conditions, the research moves closer to face detectors that function ubiquitously—across bright daylight to moonless nights, and from high-resolution DSLR photos to low-resolution security videos. This has significant implications: security systems can better detect intruders in darkness; smartphone cameras could more reliably detect faces in poorly lit scenes for autofocus or authentication; and automotive driver-monitoring systems could detect drowsiness even on dimly lit roads.

For future work, integrating temporal information (in video) or employing advanced low-light image enhancers (like learning-based methods) within this framework could further improve performance. Additionally, reducing reliance on the Viola-Jones proposals by using a light-weight learned proposal network may catch even more challenging cases (at the cost of increased complexity). The research also plans to explore the application of this hybrid model to related tasks, such as face recognition or emotion detection in low-light conditions, where it first detects the face and then passes it to recognition models. The high detection accuracy reported here provides a strong foundation to ensure that subsequent tasks have the necessary inputs. In conclusion, the hybrid CLAHE with CNN model proves to be a practical and effective solution for face detection where conventional methods fail. It underscores the importance of combining enhancement techniques with deep learning, and this research suggests that this approach can be extended to other domains that require object detection under poor visibility. By achieving robust face detection under such conditions, this work helps pave the way for more resilient vision systems that can operate in the full range of lighting and image quality conditions in the real world.

REFERENCES

[1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), vol. 1, 2001, pp. 511–518.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Adv. Neural Inf. Process. Syst. (NIPS), vol. 25, 2012, pp. 1097–1105.

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770–778.

[4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in Adv. Neural Inf. Process. Syst. (NIPS), vol. 28, 2015, pp. 91–99.

[5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 779–788.

[6] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multi-task cascaded convolutional networks," IEEE Signal Process. Lett., vol. 23, no. 10, pp. 1499–1503, 2016.

[7] J. Deng, J. Guo, Y. Zhou, and S. Zafeiriou, "RetinaFace: Single-stage dense face localisation in the wild," arXiv preprint arXiv:1905.00641, 2019.

[8] W. Yang et al., "Advancing image understanding in poor visibility environments: A collective benchmark study," IEEE Trans. Image Process., vol. 29, pp. 5737–5752, 2020.

[9] S. Yang, P. Luo, C. C. Loy, and X. Tang, "WIDER FACE: A face detection benchmark," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 5525–5533.

[10] W. Wang, W. Yang, and J. Liu, "HLA-Face: Joint high-low adaptation for low light face detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2021, pp. 16195–16204.

[11] J. Liang et al., "Recurrent exposure generation for low-light face detection," IEEE Trans. Multimedia, vol. 24, pp. 1609–1621, 2022.

[12] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, "S3FD: Single shot scale-invariant face detector," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2017, pp. 1937–1945.

[13] K. Zuiderveld, "Contrast limited adaptive histogram equalisation," in Graphics Gems IV, P. Heckbert, Ed. San Diego, CA: Academic Press, 1994, pp. 474–485.

[14] L. N. Soni, A. Datar, S. Datar "Viola-Jones algorithm based approach for face detection of African origin people and newborn infants", International Journal of Computer Trends and Technology (IJCTT) 51 (2).

[15] W. Lu, Q. Sun, and A. Li, "Improved CLAHE algorithm based on independent component analysis," in Proc. 3rd Int. Conf. Electron Inf. Technol. (EIT), IEEE, 2024, pp. 920–923.

[16] J. Li et al., "DSFD: Dual shot face detector," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2019, pp. 5060–5069.

[17] A. Neubeck and L. Van Gool, "Efficient non-maximum suppression," in Proc. 18th Int. Conf. Pattern Recognit. (ICPR), vol. 3, 2006, pp. 850–855.

[18] L. N. Soni and A. A. Waoo, "A review of recent advances in methodologies for face detection," International Journal of Current Engineering and Technology, vol. 13, no. 2, pp. 86-92, 2023.

[19] X. Yu, J. Zhang, W. Ma, and X. Zheng, "Single-stage face detection under extremely low-light conditions," in Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW), 2021, pp. 3523–3532.

[20] L. N. Soni "LNSONI Human Face Dataset", Mendeley Data, V1, (2024), doi: 10.17632/rbczppyyx8.1

[21] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.

[22] Y. P. Loh and C. S. Chan, "Getting to know low-light images with the exclusively dark dataset," Comput. Vis. Image Underst., vol. 178, pp. 30-42, 2019.

[23] L. N. Soni, A. Datar, S. Datar "Implementation of Viola-Jones Algorithm based approach for human face detection", International Journal of Current Engineering and Technology, Volume-7 Issue-5 Pp. 1819-1823.

[24] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," IEEE Trans. Image Process., vol. 26,