Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Privacy Protection: YOLOv11 Face Detection and Blurring for GDPR Compliance in Hotels

Mohammed Ikramullah khan[1], Vivekanandam B.[2]

[1]Ph. D Scholar, [2]Deputy Dean in Faculty of AI Computing and Multimedia, Lincoln University College, Malaysia

E-mail: [1]khan.ikram23@gmail.com

Abstract

Surveillance systems have undergone a drastic transformation over the years, with the advent of artificial intelligence (AI) in surveillance paving the way for better security and monitoring in public as well as private places, including hotels. But not without its considerable privacy implications since the introduction of the European Union (EU) law, the General Data Protection Regulation (GDPR), which aims to protect the privacy of EU citizens. The surveillance system collects sensitive guest data from personal information, facial data, and general appearance, making it paramount that hotels adhere to mandatory data protection laws such as the General Data Protection Regulation (GDPR) for visitors in the EU, to ensure that the data is not misused or accessed by unauthorized individuals. A privacy-protection framework for face detection and anonymization in hotel surveillance systems has been designed in this research to protect privacy from surveillance cameras based on YOLOv11, a top-tier convolutional neural network (CNN) model. The system checks for faces in video feeds/images and accurately applies a blurring mechanism, successfully anonymizing identities. The process is designed to comply with GDPR regulations while preserving essential capabilities of surveillance systems through anonymization. One of the inherent challenges is ensuring the privacy of the individuals going about their day-to-day business in front of such surveillance cameras, and at the same time, ensuring that the footage that could possibly be shared with authorities or even other stakeholders is useful. Such integration of YOLOv11 in hotel surveillance systems showcases the potential of artificial intelligence to provide security without compromising privacy.

DOI: https://doi.org/10.36548/jiip.2024.4.005
Received: 13.10.2024, received in revised form: 27.11.2024, accepted: 18.12.2024, published: 22.01.2025

Keywords: Face Detection, YOLOv11, General Data Protection Regulation (GDPR), Privacy Protection, Surveillance Systems

1. Introduction

In today's digital landscape, where technology and data collide, the hospitality industry faces an unusual conundrum—how to thread the needle between robust security and safeguarding guest privacy. Surveillance systems play a vital role in guaranteeing the security of hotels; however, they largely generate sensitive data that includes personal information such as facial identities that must be handled with caution. Increased pressure on businesses in recent years is one such factor, particularly with the introduction of the European Union's General Data Protection Regulation (GDPR), which requires hotels to comply with stringent legal standards to ensure privacy has been put in place. Artificial intelligence (AI) is a promising solution to these problems. In particular, developments in computerized eyesight and face recognition technology have opened up new opportunities for automating privacy protections.

Such progress can be seen with YOLOv11, an enhanced version of a convolutional neural network (CNN) that enables real-time detection and obfuscation of faces in video footage. This capability is especially important for sharing surveillance footage during investigations, helping to protect the identities of innocent bystanders. Implementing AI-powered solutions like these into the hotel operations not only provides compliance with regulations but also increases efficiency by helping cut and eliminate the manual labor involved in the process. At the same time, through highly automated processes which mask sensitive touchpoints, hotels can reduce errors, save precious time, as well as inspire guest trust. In this research, we examine what YOLOv11 means for the hospitality industry moving forward and how it demonstrates AI's potential to facilitate the linkage between security and compliance while redefining the bar for operational excellence.

A law on personal data processing, including facial data collected through surveillance systems, GDPR (General Data Protection Regulation) sets strict rules [25]. It is a legal framework that obliges hotels and similar businesses to block guests' disposition to prevent misuse or access of guests' data by unauthorized personnel. Addressing these problems, this study examines the use of YOLOv11, a state-of-the-art convolutional neural network (CNN),

as a mechanism for automating both face detection and blurring in real-time video. Using YOLOv11's features, hotels can blur the faces of individuals, fulfilling GDPR regulations without sacrificing safety.

Products such as YOLOv11-based machines are not only protecting privacy but also improving operational efficiency by minimizing the man-hours spent on manual tasks like video editing. Automation of face blurring minimizes human errors and time-consuming effort as a single standard system for privacy protection. Thus, implementing YOLOv11 technology aids hotels in ensuring that their security measures are compliant with the law, providing guests with peace of mind that their privacy will be respected. This approach has the potential to balance security with privacy, establishing a new standard of compliance and operational excellence in the hospitality space through AI.This approach provides an advanced surveillance solution designed to enhance railway station security through real-time tracking, missing person identification, and proactive threat responses [7].

2. Literature Review

2.1 Introduction to GDPR and Privacy Protection in Hotel

To safeguard its citizens' privacy and personal data, the European Union created the wide-ranging General Data Protection Regulation (GDPR) in May 2018. It applies to all companies (regardless of where the company is based) that process the data of EU citizens. One key piece of legislation is the General Data Protection Regulation (GDPR), considering the broad range of sensitive personal data we handle in hotels, including guest profiles, and payment details. Hotels can protect guest privacy and avoid heavy fines for noncompliance by adhering to GDPR [24].

General Data Protection Regulation (GDPR) implemented in May 2018, is a European law that protects the privacy and personal information of people residing within the EU. It applies to all companies, regardless of their location, that manage the data of European Union citizens. General Data Protection Regulation (GDPR) provides strict rules for managing personal data.

Following these rules is particularly important for hotels, which handle a lot of sensitive information from customers through surveillance techniques. GDPR requires explicit consent to data collection and enforces rights like the right to be deleted and the right to data access.

This guarantees that individuals retain authority over their personal data and can ask for its deletion or transfer; disobeying could incur hefty fines for hotels. Hotels, a sub-industry under hospitality, collect abundant personal data from their guests, such as identification documents, payment methods, and personal preferences. This data-intensive environment makes it challenging but importantly requires, GDPR compliance [11].

2.2 Evolution of CNNs in Object Detection

One of the most significant developments in the realm of computer vision (CV) is the use of Convolutional Neural Networks (CNNs) for object detection. CNNs were first proposed by [15], but became popular after successfully being applied to win the ImageNet challenge [15]. This milestone highlighted the potential of CNNs for complex visual tasks. In more recent works, research into CNN architectures has further been devoted to balancing efficiency with accuracy, as many networks are trained based on more advanced architectures [6] with the introduction of ResNet, which introduced residual learning to allow for training significantly deeper networks than were previously used.

2.3 The YOLO Series

YOLO (You Only Look Once) [17] fundamentally changed the area of object detection. The original YOLO framework unified the individual parts of object detection into a single neural network, resulting in an unprecedentedly fast detection regime with a still-competitive accuracy, an especially groundbreaking achievement for real-time applications.

YOLOv1: The idea behind YOLOv1 was to overlay an s×s grid cell over an image. If the centroid of the region of the object of interest is inside a grid space, then that grid space would detect the object. This enabled other cells to skip the item if it appeared multiple times. The YOLOv1 architecture achieved state-of-the-art accuracy on the publicly available Pascal VOC 2007 dataset with 63.4 mAP and an inference of 45 FPS (Faster Region-based Convolutional Neural Network, 2015). It should be noted that, at that time, mAP was widely accepted as the metric to assess detection performance. From a single image pass, YOLOv1 predicts the bounding boxes and class probabilities, formulating object recognition as a regression problem.

YOLOv2, or YOLO9000, in 2016, proposed the concept of fine-grained features for better small object detection and increased accuracy through a higher utilization of anchor boxes [18]. It can be trained to recognize over 9000 different object types. Anchor boxes, predefined bounding boxes, or priors were introduced in YOLOv2, serving as a set of candidate boxes that the model uses to find the best place for an object. YOLOv2 is still one of the most popular object detection techniques widely used in industrial scene maintenance and development due to its simple input and output formats, especially on low-end devices with extremely limited computing resources [21].

YOLOv3, released in 2018, was an even greater success and made a huge improvement to its robustness for various sizes of items by using three different scales for detection. More improvements were made to CNN, allowing for better feature extraction with deeper layers [18]. YOLOv3 was a significant improvement over its predecessors, achieving a mAP of 28.2 at 22 milliseconds. The YOLOv3 model relies on logistic classifiers instead of softmax and Binary Cross-entropy (BCE) loss as class prediction fundamentals.

YOLOv4, was introduced to the computer vision community in April 2020 [3], after the original inventor had moved away from the research. YOLOv4 was, in essence, the distillation of numerous proven and improved object recognition techniques into a lightweight, hard real-time object detector. YOLOv4 has tried a lot of changes, like bag-of-freebies and bag-of-specials, to find the right trade-offs. Bag-of-freebies are techniques that do not change the model during inference but alter the training method, increasing the cost [19] with data augmentation being one of the most popular technique.

YOLOv5, launched in 2020 by Ultralytics. It was an independent enhancement that prioritized usability and deploying in production contexts over the original invention [4]. This introduced changes that simplify training and improve scaling across devices. Training is easy since the model is implemented in PyTorch. The model architecture uses a Cross-stage Partial (CSP) connection hook to enhance gradient flow to reduce high computing costs. Unlike UI files, YOLOv5 uses YAML files instead of CFG files to set the model.

YOLOv6, another unofficial version, was released in 2022 by the Chinese e-commerce platform Meituan. The model was designed by the company for industrial applications and performed better than its predecessor. Compared to its predecessors, YOLOv6 features several

improvements, such as a refined network architecture for faster inference time and better handling of varying object sizes [16].

In YOLOv6 version 2.0 can minimize the accuracy drop after model quantization. Version 3.0 of YOLOv6 suggested an anchor-based head for improving accuracy to support anchor-free head learning [21]

YOLOv7, Released by a group of researchers in July 2022. They aimed to make the model faster and more accurate for the detection of items. YOLOv7 is a big milestone in the YOLO (You Only Look Once) series. Currently, YOLOv7 is an advanced object detector [21]

This evolution builds on the basic strengths of YOLO models but offers improvements tailored to real-time processing requirements and increased robustness in various operating environments. It is the fastest and most accurate object detector at 56.8% mAP at 5 to 160 frames per second. Built on Extended Efficient Layer Aggregation Network (E-ELAN), it enhances the training process by letting the model learn multiple characteristics at once through efficient computing.

YOLOv8, offers the capability to perform tasks such as instance segmentation, object identification, and image classification [9]. YOLOv8, built by Ultralytics—the same company behind the historical industry-defining YOLOv5 model—achieves high accuracy on COCO [9]. Specifically, the YOLOv8m (medium model) achieves 50.2% mAP on COCO. Against Roboflow 100, a benchmark for evaluating model performance across diverse task-specific domains, YOLOv8 surpassed YOLOv5 by a considerable margin [8]

YOLOv9, which was reported in February 2024, represents a significant breakthrough in real-time object recognition through the introduction of new techniques such as Generalized Efficient Layer Aggregation Network (GELAN) and Programmable Gradient Information [23].

The uniqueness of YOLOv9 lies in its approach to resolving information loss problems common to most deep neural networks. By integrating GELAN [23], further tunable for specific applications, with PGI, YOLOv9 achieves impressive accuracy and performance. It enhances the learning capability of the model while ensuring vital information is not lost during the detection process.

YOLOv10 introduced in May 2024, marks an essential step forward in the real-time object detection domain. This model addresses major challenges of YOLO architectures, such as reliance on NMS and computational redundancy [20], by utilizing state-of-the-art methods of YOLOv9 to boost its performance metrics and efficiency.

2.4 YOLOv11

The YOLOv11 model is the latest in the series of real-time object detection systems developed by Ultralytics, redefining parameters that enhance accuracy, speed, and operational efficiency through state-of-the-art methods [5,12] Building on the ground covered by previous versions of YOLO, the YOLOv11 framework presents significant improvements in architectural design and training techniques, making it an adaptable approach for various computer vision applications.

Key Advantages of YOLOv11

YOLOv11 shows an improvement over YOLOv9 and YOLOv10, which were launched earlier in 2024. It is superior in training techniques, feature extraction algorithms, and architectural designs [2]. Owing to its amazing speed, accuracy, and efficiency, YOLOv11 is one of Ultralytics' most powerful models yet [5,12]. Trained on data until October 2023, YOLOv11's superior framework can capture minute features better, even under non-ideal conditions. YOLOv11 introduces a range of improvements to previous versions [4]

- **More Computational Efficiency with Little or No Loss in Accuracy:** YOLOv11m has fewer parameters than YOLOv8m but is much more accurate. It outperforms YOLOv8m on the COCO dataset with 22% fewer parameters.
- **General Task Coverage:** YOLOv11 can handle a variety of CV tasks, including oriented object detection (OBB), object recognition, image classification, instance segmentation, and pose estimation.
- **Improved Speed and Efficiency:** Faster processing rates are achieved through training pipelines and more efficient architectural formulations, balancing accuracy and performance.
- **Fewer Parameters:** Models with fewer parameters are faster without significantly compromising accuracy. Enhanced Feature Extraction: YOLOv11's neck and

backbone architecture surpasses previous versions, aiding feature extraction for better detection accuracy.

- **Context Versatility:** YOLOv11 is built to be highly adaptable in various contexts, including cloud platforms, edge devices, and systems supporting NVIDIA GPUs.

The YOLO11 model variants are summarized in the below table with an emphasis on their applicability to specific tasks and their compatibility with operational modes such as Inference, Validation, Training, and Export. Because of its adaptability, YOLO11 is perfect for a wide range of computer vision applications, from complex segmentation problems to real-time detection.

Table 1. YOLOv11: Support Tasks and Models [12]

Model	Filenames	Task	Inference	Validation	Training	Export

YOLO11	`yolo11n.pt`
	`yolo11s.pt`
	`yolo11m.pt`
	`yolo11l.pt`
	`yolo11x.pt`
		Detection	✅	✅	✅	✅

YOLO11-seg	`yolo11n-seg.pt`
	`yolo11s-seg.pt`
	`yolo11m-seg.pt`
	`yolo11l-seg.pt`
	`yolo11x-seg.pt`
		Instance Segmentation	✅	✅	✅	✅

YOLO11-pose	`yolo11n-pose.pt`
	`yolo11s-pose.pt`
	`yolo11m-pose.pt`
	`yolo11l-pose.pt`
	`yolo11x-pose.pt`
		Pose/Keypoints	✅	✅	✅	✅

YOLO11-obb	`yolo11n-obb.pt`
	`yolo11s-obb.pt`
	`yolo11m-obb.pt`
	`yolo11l-obb.pt`
	`yolo11x-obb.pt`
		Oriented Detection	✅	✅	✅	✅

YOLO11-cls	`yolo11n-cls.pt`
	`yolo11s-cls.pt`
	`yolo11m-cls.pt`
	`yolo11l-cls.pt`
	`yolo11x-cls.pt`
		Classification	✅	✅	✅	✅

The below graph in Figure 1 provides the benchmark of Yolov11 against the previous versions

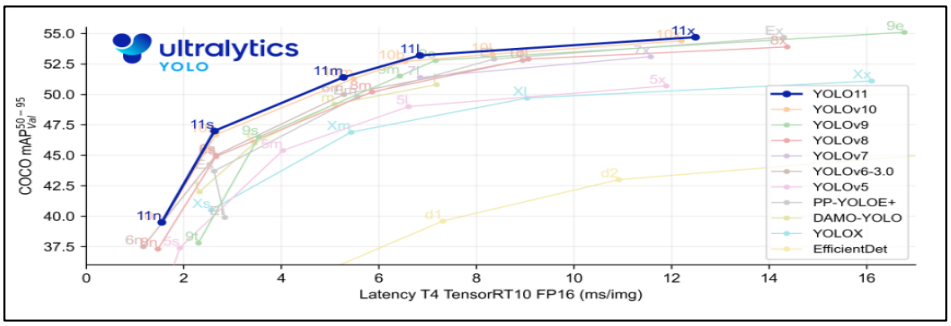

Figure 1. Benchmarking YOLOv11 Against Previous Versions [12]

The Table 1 and 2 provides the performances of YOLOv11 models on COCO dataset.

Performance Metrics

Table 2. YOLOv11 Performance Metrics [12]

Performance

Detection (COCO) | Segmentation (COCO) | Classification (ImageNet) | Pose (COCO) | OBB (DOTAv1)

See Detection Docs for usage examples with these models trained on COCO, which include 80 pre-trained classes.

Model	size (pixels)	mAP val 50-95	Speed CPU ONNX (ms)	Speed T4 TensorRT10 (ms)	params (M)	FLOPs (B)
YOLO11n	640	39.5	56.1 ± 0.8	1.5 ± 0.0	2.6	6.5
YOLO11s	640	47.0	90.0 ± 1.2	2.5 ± 0.0	9.4	21.5
YOLO11m	640	51.5	183.2 ± 2.0	4.7 ± 0.1	20.1	68.0
YOLO11l	640	53.4	238.6 ± 1.4	6.2 ± 0.1	25.3	86.9
YOLO11x	640	54.7	462.8 ± 6.7	11.3 ± 0.2	56.9	194.9

3. Methodology

3.1 Dataset

WIDER FACE is a large-scale benchmark for face detection, which is built to meet the needs of the real world. The dataset comprises 32,203 images of 13,233 labeled facial instances, presenting challenges like scale variance, pose difference, and occlusion. Details of the annotations such as occlusions, poses, and categories are well-structured and stored in an HDF5 format for the training, validation, and testing sets [24], which can even be downloaded through the links. The annotated faces include bounding boxes around important facial features and are classified by pose and occlusion levels. The evaluation protocol follows the PASCAL VOC dataset, providing two training-testing configurations. The dataset has been employed for comparing different face detection methods, exposing issues such as small scale and occlusion, which are frequent in real-world tasks like surveillance. WIDER FACE is designed to facilitate

progress in face detection research by offering a high-quality, large-scale dataset, annotated in detail to enable the evaluation of algorithm performance in challenging conditions.

3.2 YOLOv11 Architecture

YOLOv11, introduces more efficient architecture (Figure 2) with C3K2 blocks, SPFF (Spatial Pyramid Pooling Fast), and advanced attention mechanisms like C2PSA. The YOLO framework revolutionized object detection by introducing a unified neural network architecture that simultaneously handles both bounding box regression and object classification tasks, which brought a revolution in object detection. Overall, the combination of the new industry-leading features positions YOLOv11 as the optimal choice for high precision and fast processing applications, including autonomous driving and real-time surveillance systems.

Ultralytics YOLOv11 is a SOTA (state of the art) model that improves upon the previous YOLO version and adds new features and improvements to increase power and flexibility (Yolov11, 2024). It is designed with the intention of improving speed and accuracy, YOLOv11 builds on the progress made in previous versions of YOLO such as YOLOv8, YOLOv9, and YOLOv10. Built upon a better design of C3K2 blocks, SPFF (Spatial Pyramid Pooling Fast), and advanced attention mechanisms, such as C2PSA, YOLOv11 reaches a more efficient architecture. The YOLO framework revolutionized object detection by posing it as a single neural network architecture that can simultaneously perform bounding box regression and the classification of objects [13]. With its advanced capabilities, YOLOv11 stands out as an ideal candidate for a wide range of real-time applications, including autonomous driving and intelligent surveillance systems [1].

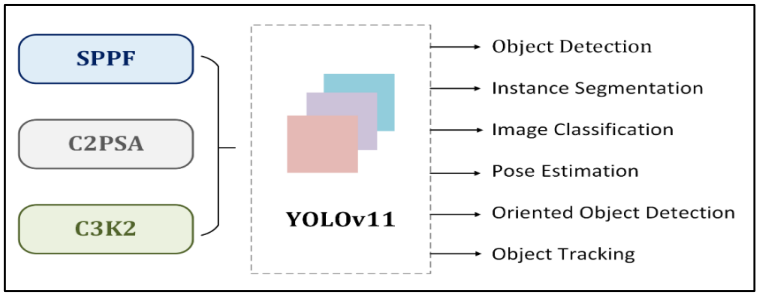

Figure 2. Key Architectural Modules in YOLOv11 [13]

3.3 Backbone

YOLOv11 backbone then extracts vital features from the input images by utilizing modern convolutional and bottleneck blocks. This allows it to form intricate patterns and well-defined details to ensure strong object detection. Through convolutional layers that gradually elevate the dimension of the channels, the backbone improves feature extraction without any superfluous computational resource consumption. The C3K2 block is integrated into the backbone design so that YOLOv11 can further improve computational performance. This block is an extension of CSP bottleneck preserving a balance between speed and accuracy by using smaller 3x3 convolution kernels, thus allowing for a more informative flow of data between layers of the neural network and better feature representation while allowing for less computational time.

3.4 Convolutional and Bottleneck Blocks

YOLOv11 follows the approach of progressively down-sampling input images with a series of convolutional blocks. By doing so, the spatial dimensions of an image are substantially reduced while the channel depth is increased. These blocks highlight semi-informed features for subsequent analysis. The older C2f block is replaced by the C3K2 block with a dramatic boost in computational efficiency and speed through feature map splitting and smaller kernel convolutions. The bottleneck block additionally helps to distill features while controlling computational load. In addition, the YOLOv11 architecture also introduces an improved version of the bottleneck block combined with residual connections inspired by ResNet to help increase gradient flow and feature pass. Such mechanisms can enhance feature representation and extraction for better learning of the network.

3.5 C3K2 Block

YOLOv11 architecture is centered on the C3K2 block. It uses small 3x3 convolution kernels that are computationally efficient and sensitive to detect prominent features. The C3K2 block enhances the information flow by fracturing feature maps and independently processing them. The C3K2 block, structurally, has a Conv block at the beginning and end, with C3K blocks in between. This design allows for a trade-off between speed and accuracy, making the C3K2 block an essential part of the YOLOv11 architecture.

3.6 C2PSA Block: Attention Mechanisms

It also introduces the C2PSA (Cross Stage Partial with Spatial Attention) block, a new attention mechanism that focuses on spatial importance in the feature map. This block emphasizes significant areas that include tiny or partially hidden objects, improving detection precision. The C2PSA block is selective focus—the layer essentially promotes the focus of specific regions to suppress noise while balancing the computational cost and accuracy.

3.7 Position-Sensitive Attention

YOLOv11's Feature Extraction with Position-Sensitive Attention passes input tensors through attention layers and concatenates the outputs to preserve positional knowledge. This mechanism improves object detection under difficult conditions like crowd scenes and overlapping images.

3.8 Detection Head

The detection head of YOLOv11 makes predictions based on multi-scale feature maps, guaranteeing a thorough detection of different object sizes. P3 is low-level features for small objects, while P5 is high-level features for larger objects. This multi-scale approach boosts YOLOv11's flexibility and strength in varied use cases. In the last step, YOLOv11 uses convolutional layers to generate bounding box coordinates and class predictions. Overall, these outputs get passed to a detect layer, which results in accurate localization and classification. YOLOv11 identifies 2x higher mAP than YOLOv8m while requiring 22% fewer parameters, with similar MACs, making it suitable for low-power devices [5]. This makes YOLO11 computationally efficient without compromising accuracy, making it suitable for deployment on resource-constrained devices.

3.9 Performance Evaluation Measures

Evaluating the performance of YOLOv11, like other object detection models, involves a combination of accuracy, efficiency, and robustness metrics. These measures assess how well the model performs across various conditions while maintaining computational efficiency. The key evaluation metrics include:

Mean Average Precision (mAP)

- The primary metric for object detection, mAP evaluates the precision-recall trade-off across multiple IoU thresholds (commonly 0.5 to 0.95).
- It is calculated as the mean of the average precision (AP) for all classes and provides a comprehensive view of model accuracy.

Precision (P) and Recall (R)

- Precision measures the proportion of correct positive detections among all predicted detections:

  P = True Positives (TP) / (True Positives (TP) + False Positives (FP))

- Recall quantifies the proportion of actual positives detected

  R = True Positives (TP) / (True Positives (TP) + False Negatives (FN))

These metrics provide a detailed understanding of YOLOv11's effectiveness and usability in real-world scenarios.

4. Results

4.1 Environmental Settings

The experimental procedures were conducted on workstation with the below specifications.

The Workstation Specifications (Figure 3) are as follows :

- Processor: Intel Core i7-8700 CPU @3.20 GHz
- RAM: 32 GB
- OS: Windows 11 Pro
- Graphic card
- CCTV Camera: Mobotix (Model: Mx-VB1A-8-IR-VA)

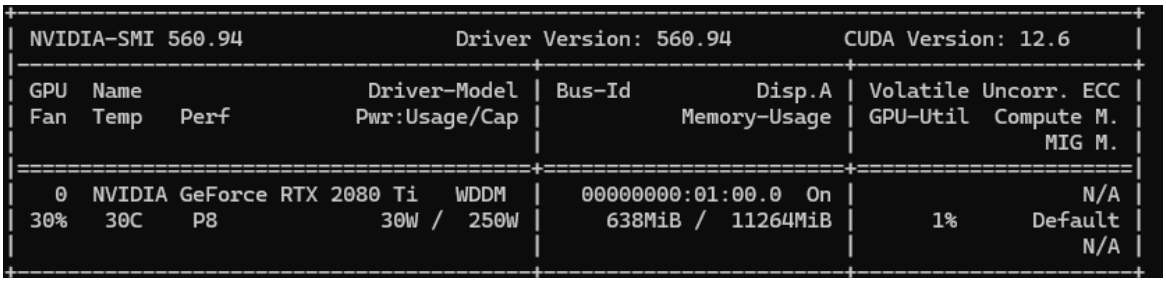

Figure 3. Workstation Specification

4.2 Experiment

The hyperparameters of the model for face detection on WIDER FACE dataset is shown below in Table. 3

The dataset was divided into training (77% - 5061 images and) and validation (23% - 1446) set. The corresponding labels were placed in labels folders.

Table 3. Hyperparameters Used

Parameters	Value
Batch-size	16
Epochs	100
Image – Size	640 x 640
Learning rate (LR)	0.0015
Momentum	0.9
Weight decay	0.0005
Optimizer	Stochastic Gradient Descent (SGD)

```yaml
# Paths
train: D:\Python\customyolo11\.venv/dataset/images/train
val: D:\Python\customyolo11\.venv\dataset\images\val

# Model parameters
nc: 1
names: ['FACE']
# Training settings
augmentation:
  mosaic: True
  mixup: 0.5
  hsv_h: 0.015  # Adjust hue
  hsv_s: 0.7    # Adjust saturation
  hsv_v: 0.4    # Adjust value
  flipud: 0.0   # Vertical flip
  fliplr: 0.5   # Horizontal flip
```

Figure 4. Yaml Coding

Table 4. Performance of YOLOv11

Model	Class	Epoch	Precision	Recall	mAP50	mAP50-95	Model - Size
YOLOv11	Face	1	0.776	0.511	0.580	0.291	109 MB
		20	0.830	0.547	0.620	0.334	
		40	0.845	0.578	0.652	0.360	
		60	0.854	0.591	0.670	0.372	
		80	0.861	0.604	0.685	0.381	
		100	0.867	0.617	0.699	0.390	

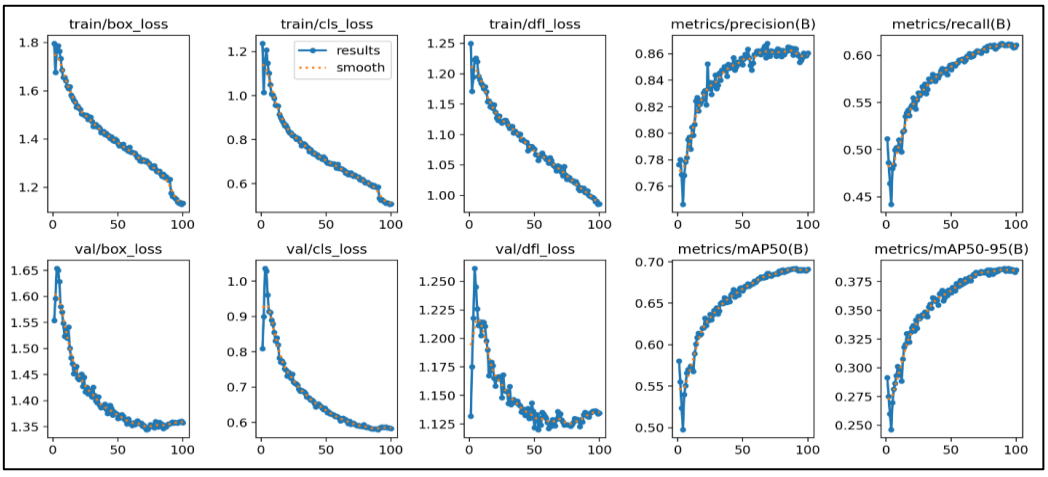

Figure 5. Performance of YOLOv11 in Training

The Figure 4 and 5, provides the coding and the visualization of various performance metrics and loss functions over the course of training a YOLOv11 model.

Below is a summary of the key trends and results derived from the graphs:

Training Loss Metrics

1. **Train/Box Loss**
    - The box regression loss decreases consistently over the training epochs, starting around 1.8 and converging to approximately 1.1. This indicates that the model is progressively improving its bounding box predictions.
2. **Train/Cls Loss (Classification Loss)**
    - The classification loss starts at around 1.2 and steadily reduces to below 0.7, suggesting that the model is becoming more accurate in identifying object classes.
3. **Train/DFL Loss (Distribution Focal Loss)**
    - This loss starts around 1.25 and decreases consistently to about 0.9, which reflects improvements in the precision of localization distributions.

Validation Loss Metrics

1. **Validation Box Loss:**
    - Similar to the training box loss, it decreases from approximately 1.65 to 1.35, confirming improved bounding box predictions on validation data.
2. **Validation Cls Loss:**
    - The validation classification loss drops from around 1.3 to 0.75, aligning well with the training loss trend.
3. **Validation DFL Loss:**
    - This loss also shows a steady reduction from around 1.25 to 1.0, indicating consistent improvement in localization accuracy for unseen data.

Performance Metrics

1. **Precision (B):**
    - Precision starts around 0.77 and improves steadily, peaking above 0.86, showing the model's increasing accuracy in minimizing false positives.

2. **Recall (B):**
    - Recall rises gradually from around 0.51 to 0.61, indicating better coverage of true positives as training progresses.
3. **mAP@50 (B):**
    - The mean Average Precision at IoU threshold 0.50 starts around 0.58 and steadily rises to 0.70, demonstrating significant improvement in overall detection accuracy.
4. **mAP@50-95 (B):**
    - The more stringent metric, mAP@50-95, shows consistent growth from 0.29 to approximately 0.39, reflecting the model's capability to perform well across varying IoU thresholds.

The YOLOv11 model demonstrates robust learning progress, with all loss metrics decreasing steadily, indicating effective optimization. Precision and recall improve consistently, and the mAP metrics confirm that the model achieves satisfactory detection accuracy. The smooth convergence of both training and validation losses also suggests the absence of significant overfitting. These results highlight the model's strong potential for real-world object detection tasks the Figure 6, 7, and8 depicts the output result observed.

Figure 6. Faces Detected and Blurred from the Image.

Figure 7. Faces Detected and Blur from the Video File.

Figure 8. Face Detected and Blur from the Live CCTV Camera

5. Conclusions

The advancements in artificial intelligence, especially in security and surveillance, offer valuable solutions for the hospitality industry, where ensuring guest safety and privacy is critical. This article explores how YOLOv11, a powerful AI model, can address these challenges by detecting and anonymizing faces in real-time. This not only protects sensitive data, complying with GDPR, but also saves time by automating the process and reducing the chances of human error. By adopting such technologies, hotels can improve their security practices while maintaining guest trust and privacy. YOLOv11 demonstrates how AI can effectively balance security needs with privacy requirements, setting a new standard for efficiency and compliance in the hospitality sector.

The current study utilized a limited training setup, which restricted the model's full potential. Future research could focus on training YOLOv11 on more extensive datasets, using

higher-end hardware, and running for more epochs to further enhance its accuracy and performance. These improvements could yield even better results, solidifying the role of AI in addressing the complex demands of security and privacy in the hospitality industry.

References

[1] Alif, M. A. R. (2024). YOLOv11 for Vehicle Detection: Advancements, Performance, and Applications in Intelligent Transportation Systems (No. arXiv:2410.22898). arXiv. https://doi.org/10.48550/arXiv.2410.22898

[2] Ankan Ghosh. (2024). YOLO11: Faster Than You Can Imagine! https://learnopencv.com/yolo11/

[3] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection (No. arXiv:2004.10934). arXiv. https://doi.org/10.48550/arXiv.2004.10934

[4] Glenn Jocher. (2020). YOLOv5. https://github.com/ultralytics/yolov5

[5] Glenn Jocher, & Jing Qiu. (n.d.). YOLO11 🚀 NEW. Https://Docs.Ultralytics.Com/Models/Yolo11/. Retrieved December 8, 2024, from https://docs.ultralytics.com/models/yolo11

[6] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition (No. arXiv:1512.03385). arXiv. https://doi.org/10.48550/arXiv.1512.03385

[7] Jyothis, Siva, Anand S. Nair, Kirandas VR, and Divya Visakh. "Developing a SmartRail Security System with YOLO and OpenCV." Journal of Ubiquitous Computing and Communication Technologies 6, no. 1 (2024): 28-38.

[8] Hussain, M. (2024). YOLOv5, YOLOv8 and YOLOv10: The Go-To Detectors for Real-time Vision (No. arXiv:2407.02988). arXiv. https://doi.org/10.48550/arXiv.2407.02988

[9] Jacob Solawetz & Francesco. (2024, October 23). What is YOLOv8? A Complete Guide. Roboflow Blog. https://blog.roboflow.com/what-is-yolov8/

[10] Jane Austen. (2024, February 10). What is New in YOLOv8; Deep Dive into its Innovations-Yolov8. https://yolov8.org/what-is-new-in-yolov8/

[11] Jhuang, Yun-Yun, Yu-Hui Yan, and Gwo-Jiun Horng. "GDPR Personal Privacy Security Mechanism for Smart Home System." Electronics 12, no. 4 (2023): 831.

[12] Jocher, G., & Qiu, J. (2024). Ultralytics YOLO11 (Version 11.0.0) [Computer software]. https://github.com/ultralytics/ultralytics

[13] Khanam, R., & Hussain, M. (2024). YOLOv11: An Overview of the Key Architectural Enhancements (No. arXiv:2410.17725). arXiv. https://doi.org/10.48550/arXiv.2410.17725

[14] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Advances in neural information processing systems 25, 1097–1105.

[15] Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

[16] Li, C., Li, L., Geng, Y., Jiang, H., Cheng, M., Zhang, B., Ke, Z., Xu, X., & Chu, X. (2023). YOLOv6 v3.0: A Full-Scale Reloading (No. arXiv:2301.05586). arXiv. https://doi.org/10.48550/arXiv.2301.05586

[17] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection (No. arXiv:1506.02640). arXiv. https://doi.org/10.48550/arXiv.1506.02640

[18] Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, Faster, Stronger. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 6517–6525. https://doi.org/10.1109/CVPR.2017.690

[19] Terven, J., & Cordova-Esparza, D. (2023). A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS.

Machine Learning and Knowledge Extraction, 5(4), 1680–1716. https://doi.org/10.3390/make5040083

[20] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). YOLOv10: Real-Time End-to-End Object Detection (No. arXiv:2405.14458). arXiv. https://doi.org/10.48550/arXiv.2405.14458

[21] Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors (No. arXiv:2207.02696). arXiv. https://doi.org/10.48550/arXiv.2207.02696

[22] Wang, C.-Y., & Liao, H.-Y. M. (2024). YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems (No. arXiv:2408.09332). arXiv. https://doi.org/10.48550/arXiv.2408.09332

[23] Wang, C.-Y., Yeh, I.-H., & Liao, H.-Y. M. (2024). YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information (No. arXiv:2402.13616). arXiv. https://doi.org/10.48550/arXiv.2402.13616

[24] Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). WIDER FACE: A Face Detection Benchmark. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). http://shuoyang1213.me/WIDERFACE/

[25] Zhang, L., Moukafih, N., Alamri, H., Epiphaniou, G., & Maple, C. (2024). A BERT-based Empirical Study of Privacy Policies' Compliance with GDPR (No. arXiv:2407.06778). arXiv. https://doi.org/10.48550/arXiv.2407.06778