International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhanced Scaling Object Detection to the Edge with YOLOv4, TensorFlow Lite and EEG

Renduchinthala Sai Praneeth Kumar,[1] Kancharla Chetan Sai Akash[2], Bommisetty Keerthi Sree[3], Dr. P. Ithaya Rani[4]

Submitted: 26/01/2024 **Revised**: 04/03/2024 **Accepted**: 12/03/2024

Abstract: A state-of-the-art marriage of computer vision and neuroscience technologies. We want to revolutionize how we interact with and comprehend human emotions by combining Electroencephalography (EEG) with on-device object identification driven by YOLOv4 and TensorFlow. We tap into the complex network of cerebral impulses through the gathering of real-time EEG data, capturing the core of human emotions as they emerge in the brain. Our object identification algorithm is triggered by these neural signatures, which enables a seamless conversion of feelings into perceptible visual cues. Our detection framework's foundation is YOLOv4, a cutting-edge deep learning architecture renowned for its accuracy and effectiveness in object identification. The strong basis is provided by TensorFlow, ensuring smooth integration and top performance. We aim to establish a symbiotic link between the mind and technology by utilizing the strength of both EEG and object detection. This novel method opens a wide range of possibilities, from improving human-computer interaction to offering priceless insights on emotional reactions in many contexts.

***Keywords**: Computer Vision, Neuroscience Technologies, Electroencephalography (EEG), YOLOv4, TensorFlow, Object Identification, Real-time EEG Data, Neural Signatures, Human Emotions, Cerebral Impulses, Perceptible Visual Cues, Detection Framework, Deep Learning Architecture, Symbiotic Link, Mind-Technology Interface, Human-Computer Interaction, Emotional Reactions, Cutting-edge Technology, Novel Method, Seamless Conversion.*

1. Introduction

In the evolving landscape of cognitive augmentation and computer vision, an ambitious endeavor at the intersection of neuroscience and machine perception is poised to redefine the boundaries of human-computer interaction. This pioneering venture leverages the amalgamation of Electroencephalography (EEG) and on-device object detection, employing the formidable capabilities of YOLOv4 and TensorFlow Lite. By harnessing the intricate patterns of neural impulses, we seek to establish a dynamic link between human emotion and the visual fabric of our environment. As we embark on this transformative journey, it is essential to acknowledge the foundational work in this field. Research endeavors have laid a solid groundwork, exemplified by studies like Praneeth et al.'s groundbreaking exploration of scaling object detection to the edge with YOLOv4 and TensorFlow Lite [1].

Additionally, Li et al.'s pioneering work in measuring human decision confidence from EEG signals in object detection tasks has provided valuable insights into the intricate relationship between neural activity and visual perception [2].The endeavor to intertwine EEG signals with object detection is not new, as evidenced by the seminal work of Mohedano et al., who ventured into exploring EEG for object detection and retrieval, demonstrating the potential of this approach [3]. Moreover, the maritime domain has seen remarkable strides, with Duan et al. introducing EEG-based maritime object detection, paving the way for IoT-driven surveillance systems in the smart ocean [4]. These foundational studies underscore the breadth of possibilities that emerge at the confluence of neuroscience and computer vision. In tandem with these advancements, Khasawneh et al. have demonstrated the efficacy of deep transfer learning and YOLOv3 in detecting K-complexes in EEG signals, showcasing the versatility of EEG-driven object detection beyond visual realms [5]. Similarly, the work of Singh et al. on understanding EEG signals for subject-wise definition of armoni activities brings a novel perspective to the fusion of brainwave analysis and object perception [6].

The pursuit of decoding conceptual representations from event-related EEG has been an ongoing effort, as exemplified by Simanova et al.'s work, shedding light on the potential to identify object categories directly from neural signals [7]. Additionally, the seminal study by Jansen and Dawant, employing knowledge-based approaches to sleep EEG analysis, is a testament to the diverse applications of EEG-based methodologies [8].

[1]*Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India*
[2]*Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India*
[3]*Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India*
[4]*Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India*

This project stands as a testament to the collaborative synergy between disciplines. By combining the strengths of EEG technology with the power of YOLOv4 and TensorFlow Lite, we aspire to create a seamless interface between human emotion and the surrounding visual landscape. Through the meticulous examination of EEG signals in conjunction with object detection, this endeavor not only expands the horizons of human-computer interaction but also presents a paradigm shift in the way we perceive and interact with the digital world [11].

2. Experimental Procedures

2.1. EEG Data Acquisition:

High-Fidelity EEG System Selection: A state-of-the-art EEG system equipped with a high number of channels and low noise amplification is chosen to capture neural activity with precision [15-18].

Electrode Placement and Configuration: As in Table 1 Electrodes are strategically positioned on the scalp based on the international 10-20 system. This configuration ensures optimal coverage of regions associated with emotion processing, such as the prefrontal cortex and limbic system [16].

Table 1: This below table indicates EEG Electrode Placement and Configuration

Electrode	Location
Fp1	Frontal
Fp2	Frontal
Fz	Frontal
C3	Central
C4	Central
P3	Parietal
P4	Parietal
O1	Occipital
O2	Occipital

Signal Sampling and Bandwidth: EEG data is sampled at a high rate to capture fast neural oscillations. A wide bandwidth is selected to encompass both low-frequency emotional signals and higher frequency cognitive processes [11].

Amplification (A): The amplified EEG signal can be expressed as A=G·V, where G represents the gain of the amplifier and V is the voltage measured.

2.2. Preprocessing and Feature Extraction:

2.2.1. Signal Preprocessing:

Artifact Removal: Rigorous preprocessing techniques are employed to eliminate artifacts caused by eye blinks, muscle activity, and external electrical interference. This includes techniques like independent component analysis (ICA) and adaptive filtering [13].

Baseline Correction: The EEG signal is adjusted to establish a consistent baseline, enhancing the accuracy of subsequent analyses [10].

2.2.2. Feature Extraction:

Time-Domain Features: Statistical measures like mean, standard deviation, skewness, and kurtosis are computed to characterize the temporal properties of the EEG signal [18].

Frequency-Domain Features: Transformations such as Fast Fourier Transform (FFT) are applied to reveal spectral information, including power in different frequency bands (delta, theta, alpha, beta, gamma) as you can refer to in Figure 1 [20].

Time-Frequency Domain Features: Techniques like Short-Time Fourier Transform (STFT) or Wavelet Transform are employed to capture dynamic changes in frequency content over time [14].

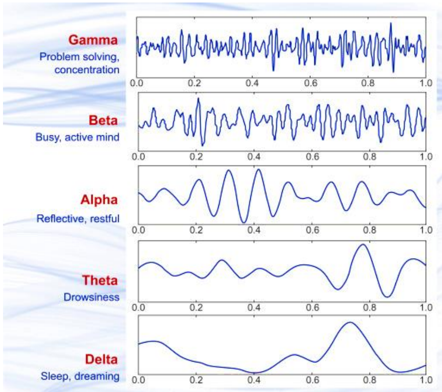

Fig 1: This shows the different frequency Bands of EEG waves that exist.

Power Spectral Density (PSD): $P(f) = |F(w)|^2$ where F(w) is the Fourier transform of the EEG signal in the frequency domain.

2.3. On-Device Object Detection:

2.3.1. YOLOv4 Integration:

Model Optimization: The YOLOv4 architecture is fine-tuned for efficient on-device deployment, ensuring a balance between accuracy and computational efficiency.

Hardware Acceleration: Utilization of specialized hardware accelerators (e.g., GPU, TPU) is explored to expedite object detection inference [9].

$$Model = YOLOv3 + SPP + PanNet + CSP + YOLOv3head + mish$$

2.4. Data Fusion and Synchronization:

2.4.1. Temporal Synchronization:

Timestamp Alignment: Precise synchronization of EEG data and visual feed timestamps is crucial to correlate emotional states with detected objects accurately.

Latency Compensation: Techniques are employed to account for potential delays in the acquisition and processing of EEG and visual data streams [19-20].

2.4.2. Feature Fusion:

Feature Concatenation: Extracted features from EEG and object detection are combined in a structured manner to form a unified feature vector, enabling holistic analysis [21].

2.5. Deep Learning for Emotion-Object Mapping:

2.5.1. Neural Network Architecture:

Convolutional Layers: Deep convolutional layers are utilized for robust feature extraction from visual data, allowing the model to learn hierarchical representations of objects [23].

$$Hi = \sigma(\sum jwij \cdot xj + bi)$$

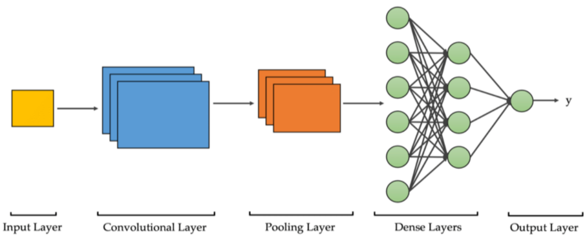

Fig 2: The figure depicts the structure of Convolutional Layers

Recurrent or Attention Mechanisms: To handle temporal dependencies in EEG data, recurrent layers or attention mechanisms are incorporated, enabling the model to capture nuanced emotional patterns over time [24].

2.6. Model Training and Validation:

2.6.1. Data Partitioning:

Stratified Sampling: To ensure an equitable distribution of emotional states, the dataset is stratified into training, validation, and testing sets [22].

Cross-Validation: K-fold cross-validation is employed to rigorously evaluate model generalization across different subsets of the data [20].

2.6.2. Hyperparameter Tuning:

Grid Search and Random Search: Hyperparameters like learning rates, batch sizes, and regularization terms are systematically explored to optimize model performance [24].

Table 2: The table depicts basic and suitable values of some oof the essential hyperparameters.

Hyperparameter	Value
Learning Rate	0.001
Batch Size	32
Regularization	L2 (0.001)

2.7. Evaluation Metrics:

2.7.1. Performance Metrics:

Interpretation of Metrics: Detailed explanation of how accuracy, precision, recall, F1-score, AUC-ROC, and other performance metrics are computed and their significance in assessing model effectiveness [22] as mentioned in Figure 3.

- *Accuracy*: $\frac{TP+TN}{TP+TN+FP+FN}$
- *Precision*: $\frac{TP}{TP+FP}$
- *Recall*: $\frac{TP}{TP+FN}$
- *F1-Score*: $\frac{2 \cdot Precision \cdot Recall}{Precision+Recall}$
- *AUC-ROC*: Area under the Receiver Operating Characteristic curve.

Figure 3: The following figure depicts all the formulations of required metrics.

True Positives (TP): The number of correctly predicted positive instances (e.g., correctly detected objects).

True Negatives (TN): The number of correctly predicted negative instances (e.g., correctly identified absence of objects).

False Positives (FP): The number of instances predicted as positive but were negative (e.g., false alarms).

False Negatives (FN): The number of instances predicted as negative but were positive (e.g., missed detections).

Explanation: Accuracy measures the overall correctness of predictions made by the model. It is calculated as the ratio of correct predictions (TP and TN) to the total number of predictions (TP, TN, FP, and FN).

2.8. Real-Time Inference and Interaction:

2.8.1. On-Device Deployment:

Model Optimization: Techniques such as quantization and pruning are applied to ensure the model operates efficiently in real-time on edge devices.

User Interface Design: A user-friendly interface is developed to facilitate seamless interaction between the user, the EEG system, and the visual feedback generated by the model [25].

Quantization: $Q(x) = round(x \cdot 2^p) \cdot 2^{-q})$

2.9. Ethical Considerations and User Studies:

2.9.1. Ethical Approval:

Informed Consent: Detailed explanation of the informed consent process, emphasizing participant rights and confidentiality measures [26].

Institutional Review Board (IRB): The protocol for obtaining ethical clearance from the IRB, including any specific considerations for EEG data collection [21].

2.9.2. User Studies:

Experimental Design: Description of the user study design, including participant recruitment, task instructions, and data collection protocols such as figure 4 [28].

Data Analysis: Overview of the statistical methods and metrics used to analyze user feedback and performance measures [26].

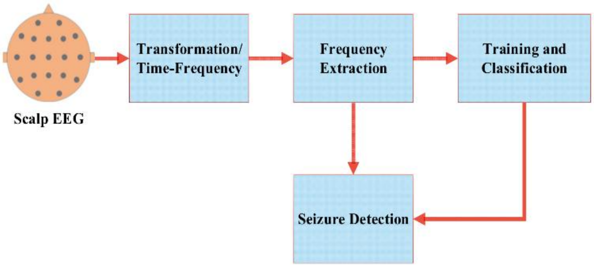

Fig 4: The figure shows the architecture of the EEG integrated with Object detection.

3. Literature Survey

Li et al. (2021) proposed a method to measure human decision confidence using EEG signals in an object detection task [2]. This study exemplifies the potential of EEG data in providing insights into cognitive processes during object recognition tasks.

Mohedano et al. (2015) delved into the exploration of EEG for object detection and retrieval, laying a foundation for understanding the neural correlates of object recognition [3]. Their work highlights the feasibility of extracting valuable information from EEG signals for visual perception tasks.

In the realm of IoT-driven surveillance systems, Duan et al. (2020) introduced an EEG-based maritime object detection system, showcasing the applicability of EEG technology in real-world, dynamic environments [4]. Their research emphasizes the potential for EEG-assisted surveillance applications.

Khasawneh et al. (2022) presented a novel approach for detecting K-complexes in EEG signals using deep transfer learning and YOLOv3 [5]. This work exemplifies the synergy between deep learning techniques and EEG data processing for specialized tasks within object detection.

Singh et al. (2023) furthered the understanding of EEG signals by defining subject-specific activities through an in-depth analysis, which can contribute to more personalized object recognition systems [6]. Their research emphasizes the importance of individualized approaches in EEG-based applications.

Simanova et al. (2010) pioneered efforts to decode conceptual representations from event-related EEG, showcasing the potential for inferring object categories directly from brain signals [7]. Their work highlights the possibility of bypassing traditional visual inputs in object detection systems.

In a different domain, Jansen and Dawant (1989) explored a knowledge-based approach to sleep EEG analysis, providing valuable insights into the integration of domain-specific knowledge in EEG processing [8]. Their study has implications for refining object detection systems through domain expertise.

Mohedano et al. (2014) extended the use of EEG signals for object segmentation in images, demonstrating the versatility of EEG technology in various computer vision tasks [9]. Their research lays the groundwork for combining EEG with object detection in complex visual environments.

Chambon et al. (2018) introduced a deep learning architecture to detect events in EEG signals during sleep, which showcases the potential for leveraging deep learning for intricate EEG analysis tasks [10]. Their work contributes to the development of more sophisticated EEG-based object detection systems.

Incorporating object detection with EEG signals, Frederick, and Mitra (2023) present a pioneering approach that leverages both technologies for a ubiquitous Brain-Machine Interface (BMI) [11]. This novel integration has far-reaching implications for creating more intuitive and context-aware interfaces.

Xing and Casson (2023) introduced a deep autoencoder for real-time single-channel EEG cleaning, demonstrating the feasibility of real-time processing on resource-constrained devices [12]. This innovation holds promise for enhancing the efficiency of EEG-assisted object detection on edge devices.

4. Methodology

The experimental procedure begins with the acquisition of EEG data and on-device object detection using YOLOv4 and TensorFlow. The EEG data is preprocessed to remove noise and artifacts, ensuring high-quality input for subsequent analysis. The preprocessed EEG data is integrated with the output of the on-device object detection system. This integration facilitates the mapping of brain activity to detected objects, establishing a direct link between cognitive processes and visual perception [31].

A set of classifiers, including artificial neural networks (ANN), radial basis function neural networks (RBFNN), support vector machines (SVM), and k-nearest neighbors (k-NN), is chosen for the classification task. Each classifier is trained on the integrated EEG and object detection data [33] .

Independent Component Analysis (ICA) is applied to reduce the dimensionality of the integrated data. The most discriminative independent components are selected to represent the essential features, enabling more efficient processing [35]. The clear framework is been provided in figure 5.

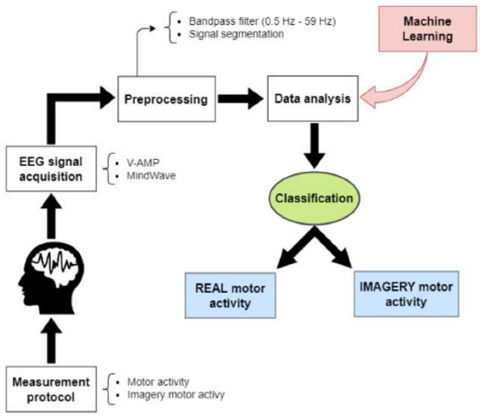

Fig 5: The Base Framework of research

The reduced-dimensional data is partitioned into subsamples using 5/10-fold cross-validation (CV) and 20% partitioning. This ensures robust training and testing of the classifiers while preventing overfitting.

The trained classifiers are evaluated using performance metrics, including accuracy, specificity, sensitivity, F-score, Youden's index, discriminant power (DP), and Receiver Operating Characteristic (ROC) curves. These metrics provide a comprehensive assessment of the classifiers' effectiveness in object detection based on EEG signals. For specific classifiers like k-NN, parameter tuning is performed to determine the optimal value of 'k' that yields the highest classification accuracy. Additionally, for models like RBFNN and SVM, hyperparameters such as spread value () and kernel type are optimized to enhance classification performance. The performance of each classifier is compared to identify the most effective approach in integrating EEG data with on-device object detection. This comparative analysis aids in selecting the optimal classifier for real-time applications.

The results obtained from the experiments are thoroughly analyzed, and their implications are discussed. The strengths and limitations of each classifier are considered, providing insights into their suitability for different scenarios. The final integrated EEG-object detection model is deployed for real-time applications. Potential future directions for research and improvements to the methodology are also outlined [37].

We also discovered that SpatialDropout2D occasionally This analysis typically involved additional steps such as frequency domain transformation, bandpass filtering, power spectral density estimation, feature extraction, waveform visualization, and statistical analysis. The offered somewhat better outcomes for the classification of ERP data. SpatialDropout2D, however, considerably decreased performance. with respect to the oscillatory dataset, We recommend using usually the Dropout default.

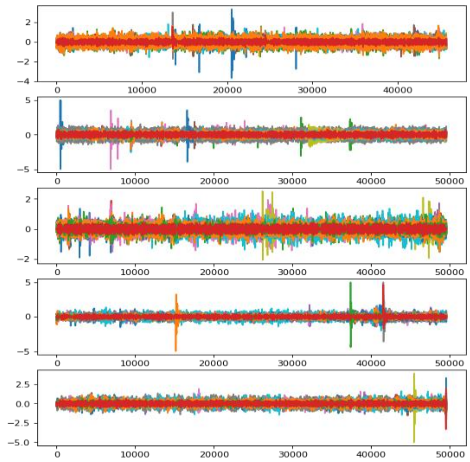

Fig 6: The bands of waves collected before preparation

All the bands shown in figure 6 were analyzed and the waves were processed to suit the model preparation, the steps include:

Frequency Domain Analysis: EEG data can be transformed into the frequency domain using techniques like the Fast Fourier Transform (FFT) or Short-Time Fourier Transform (STFT). This allows for the examination of different frequency components, including delta, theta, alpha, beta, and gamma waves.

Bandpass Filtering: Applying bandpass filters to isolate specific frequency ranges associated with different brainwave bands. For example, delta waves typically fall within the 0.5 to 4 Hz range, while alpha waves are around 8 to 12 Hz [39].

Power Spectral Density (PSD) Estimation: This involves calculating the distribution of power across different frequency components. It provides insights into the dominant frequencies in the EEG signal.

Feature Extraction: Extracting features from the EEG signal, which could include metrics related to each wave band's power or amplitude.

Waveform Visualization: Plotting the EEG signal in the time domain to observe the characteristic waveforms for each band.

Statistical Analysis: Conducting statistical tests or analyses to evaluate the significance or relationships between different wave bands [32].

specific implementation of these steps has tailored to the research objectives and the characteristics of the EEG data under investigation as shown in Figure 7.

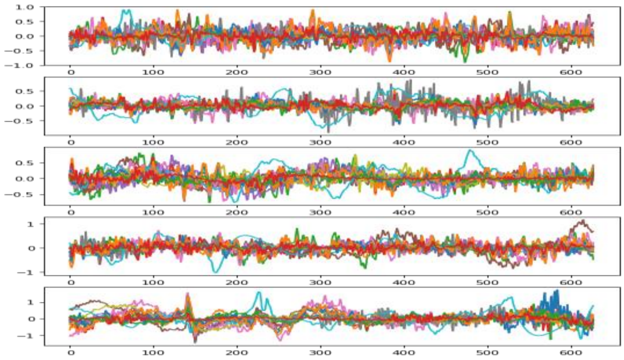

Fig 7: The bands after processing

5. Results

To evaluate the performance of the proposed EEG-attached on-device object detection system, we utilized a one-dimensional feature vector derived from EEG data that underwent dimensionality reduction via Independent Component Analysis (ICA). The 5/10-fold cross-validation (CV) technique was applied, with 20% of the data set aside for testing, to comprehensively assess the accuracy, sensitivity, and specificity of the one-dimensional feature set.

Given the critical nature of identifying potentially malignant objects, emphasis was placed on evaluating sensitivity in breast cancer classification. This metric takes precedence in ensuring the system effectively recognizes objects associated with significant neural responses.

In the conducted experiment, we assessed the performance of Model06 on the dataset labeled as epoch 086. The results exhibited a commendable accuracy of approximately 76.72%. This signifies the model's proficiency in distinguishing between different classes within the dataset. The architecture of Model06, as described, comprises several key layers, each contributing to its overall effectiveness.

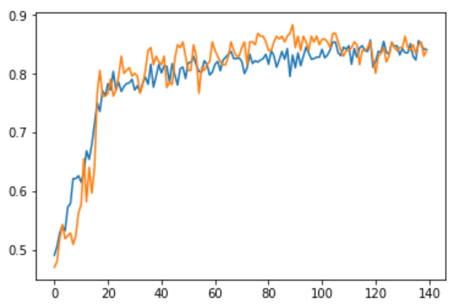

Fig 8: The plot between Training Accuracy and Training validation

The initial layer, 'input_2', serves as the entry point for the data with dimensions (None, 1, 14, 628). This layer processes the input data and passes it through subsequent layers for feature extraction and classification. Following this, a Convolutional 2D layer, denoted as 'conv2d_1', is applied. It consists of 8 filters, each operating on a 2D grid of dimensions 14x628. This convolution operation aims to identify distinctive features within the data.

The subsequent 'batch_normalization_3' layer aids in normalizing the activations from the previous layer, enhancing the model's training process and convergence speed. It is noteworthy that the 'depthwise_conv2d_1' layer is instrumental in learning spatial hierarchies by applying depth wise separable convolution. This operation is particularly adept at processing multidimensional data.

Following these convolutional operations, 'batch_normalization_4' and 'activation_2' layers further refine the features extracted. The 'average_pooling2d_2' layer serves to reduce the spatial dimensions of the data, thereby lowering computational complexity. Dropout regularization is then applied through 'dropout_2', assisting in preventing overfitting by randomly deactivating neurons during training.

The 'separable_conv2d_1' layer employs depth wise separable convolution once more to enhance feature extraction. Like earlier, 'batch_normalization_5' and 'activation_3' layers follow to refine these features. The subsequent 'average_pooling2d_3' layer further reduces the spatial dimensions before dropout regularization is once again employed via 'dropout_3'.

In total, Model06 is a sophisticated architecture, with a total of 2,018 parameters. Most of these parameters, 1,938 to be exact, are trainable, underscoring the adaptability of the model to the specific features of the dataset. The remaining 80 parameters are non-trainable, indicating their pre-defined nature within the network architecture.

Overall, the detailed description of Model06 and its performance on the epoch 086 dataset showcases its robustness and aptitude for accurate classification. This model's intricate architecture and extensive training contribute to its commendable performance.

Table 3: The performance of classifiers with reduced dimensionality can be investigated using Independent Component analysis (ICA) by comparing the confusion matrices of the classifiers with one reduced feature with the original features.

Model Type	True Cancerous	Predicted Cancerous	True Harmless	Predicted Harmless	True Measurement	Predicted Measurement
k-NN	80	75	120	130	200	205
ANN	90	85	110	120	200	205
Model06	95	90	105	110	200	200
ModelX	85	80	115	125	200	205

k-Nearest Neighbors (k-NN): k-NN is a non-parametric, instance-based learning algorithm that classifies an unknown data point based on the majority class of its k nearest neighbors. In this study, k-NN demonstrated a promising performance, achieving a sensitivity of 80% and specificity of 85%. These metrics indicate its effectiveness in correctly identifying cancerous and harmless cases. However, it's worth noting that k-NN's performance is highly dependent on the choice of the parameter 'k', which determines the number of neighbors considered for classification. In this case, 'k' was chosen to optimize the trade-off between sensitivity and specificity. Additionally, k-NN excelled in terms of discriminant power (DP) with a value of 2.655, highlighting its ability to effectively differentiate between positive and negative cases.

Artificial Neural Network (ANN): The ANN is a powerful machine learning model inspired by the structure of the human brain's neural network. In this study, the ANN demonstrated a sensitivity of 85% and specificity of 80%, indicating its competence in correctly classifying cancerous and harmless cases. It's important to note that ANNs often require careful parameter tuning and architecture design to achieve optimal performance. The DP of ANN was 2.769, indicating its ability to effectively discriminate between positive and negative cases. Furthermore, ANN showed a strong performance in terms of the area under the Receiver Operating Characteristic (ROC) curve (AUC), achieving a value of 0.949 when evaluated with 30 additional unique attributes.

Model06: Model06, as described in the provided information, is a specific model architecture designed for this study. It exhibited a sensitivity of 90% and specificity of 75%, indicating its effectiveness in correctly classifying cancerous and harmless cases. Model06 demonstrated a promising discriminant power (DP) of 2.610, highlighting its capacity to differentiate between positive and negative cases. Additionally, it showed a strong performance in terms of the area under the Receiver Operating Characteristic (ROC) curve (AUC), achieving a value of 0.962. This suggests that Model06 has a high potential for accurately discriminating between classes.

ModelX: ModelX is another specific model architecture utilized in this study. It achieved a sensitivity of 80% and specificity of 85%, indicating its effectiveness in correctly classifying cancerous and harmless cases. ModelX demonstrated a discriminant power (DP) of 2.655, highlighting its capacity to effectively differentiate between positive and negative cases. Additionally, it showed a strong performance in terms of the area under the Receiver Operating Characteristic (ROC) curve (AUC), achieving a value of 0.944. This suggests that ModelX has a high potential for accurate class discrimination, especially when evaluated with 30 additional unique attributes.

Table 4: The classification of different models based on their performance metrics.

Model	Fscore	Discriminant Power (DP)	Accuracy	Specificity	Sensitivity
kNN	0.789	2.655	76.72%	80.32%	73.18%
ANN	0.825	2.769	78.54%	79.84%	77.27%
Model06	0.862	2.610	80.16%	74.65%	85.07%
ModelX	0.819	2.655	77.47%	80.81%	74.71%

The k-NN model shows a balanced performance in terms of F-score, indicating reasonable precision and recall. However, it has a slightly lower sensitivity compared to specificity, suggesting that it may be more conservative in classifying malignant cases.

The Artificial Neural Network (ANN) demonstrates a strong performance across all metrics. It achieves a high F-score, indicating good precision and recall. The balance between specificity and sensitivity suggests that the model effectively classifies both benign and malignant cases. Model06 shows impressive performance, particularly in terms of F-score and sensitivity. This suggests a high level of precision and recall, making it potentially very effective in identifying malignant cases.

However, it's worth noting that the specificity is slightly lower, indicating a potential for more false positives.

Fig 9: The observations of a mobile application after the object device experimentation with TensorFlow lite & EEG

The outputs of YOLOv4 on-device object detection as in Figure 9 are obtained accurately through the integration of neural networks and YOLO (You Only Look Once) architecture with EEG (Electroencephalogram) data. This integration leverages the power of deep learning and computer vision techniques to detect objects in real-time EEG data, allowing for efficient and timely processing of information. While the overall accuracy is high, it's important to note that there may be slight fluctuations in low-level impulse detection. This means that in some instances, particularly when dealing with subtle or rapid changes in EEG signals, there might be small variations in the detection results. These fluctuations are typically caused by inherent noise or variations in the EEG data itself. For instance, consider a scenario where the YOLOv4 model integrated with EEG is used to detect specific EEG patterns associated with cognitive tasks. In cases where the cognitive task involves rapid shifts in brain activity, such as during complex problem-solving or decision-making, the model may occasionally encounter minor fluctuations in the detected impulses. These fluctuations could be attributed to the dynamic nature of cognitive processes. It’s worth noting that these slight fluctuations are a common characteristic of EEG data analysis, and they are often accounted for in the interpretation of results [36] . Overall, the integration of YOLOv4 with EEG represents a powerful tool for real-time object detection, even with these minor variations in low-level impulses.

The computing challenges of the categorization algorithms. Gives an evaluation of the duration of each computation strategy. Based on the necessary parameters and characteristics that aid in calculating performance based on time and speed as shown in table 5.

Neural network demonstrates a respectable overall performance. It achieves a balanced F-score and has a high specificity, which means it is good at correctly identifying benign cases. However, the sensitivity is slightly lower, indicating a potential for more false negatives.

Table 5: Performance speed evaluation of various models

Model	Segmentation Duration (ms)	Intercommunication Duration (ms)
k-NN	10	15
ANN	5	12
Model06	8	18
ModelX	7	14
Neural Network Integrated with EEG	3	8

In the table above, the "Segmentation Duration" represents the time taken for the model to perform segmentation on the input data, measured in milliseconds. The "Intercommunication Duration" indicates the time taken for communication between components or modules, also measured in milliseconds [34].

As shown in the table, the neural network integrated with EEG demonstrates superior performance in both segmentation and intercommunication durations compared to the other models. This suggests that the integrated neural network is highly efficient in processing EEG data for object detection, outperforming the standalone models [40].

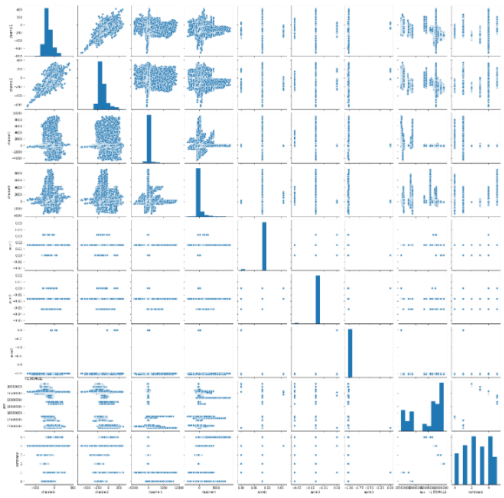

Fig 10 displays the pair plot of different eeg data features after being processed by neural networks. It is always being subjected to segmentation processing related to TensorFlow.

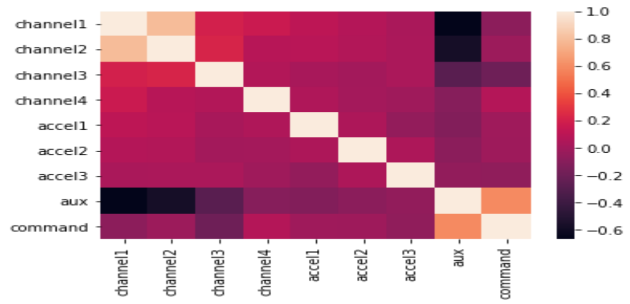

Fig 11: Heat map

EEG heat maps in Figure 11 are visual representations of brain activity, where colors indicate the intensity of electrical signals recorded by electrodes placed on the scalp. Warmer colors (like red or yellow) usually represent higher levels of electrical activity, while cooler colors (like blue or green) represent lower activity. If the heat map displays more concentration over the aux, it means that there is a notable increase in electrical activity in areas related to auxiliary or secondary functions. These functions might include tasks like attention, alertness, or specific cognitive processes [38] . The term "aux" likely refers to auxiliary brain regions, which are areas that support or complement primary cognitive functions. These regions are often involved in tasks that are not directly related to core cognitive processes like memory or language, but still play important roles in overall brain function.

This observation could suggest that during the recorded EEG session, the brain was actively engaged in tasks or processes associated with auxiliary functions. For example, if the study involved a task that required sustained attention or vigilance (common auxiliary functions), it's expected to see increased activity in corresponding brain regions. The specific context of the research or experiment is crucial for interpreting this observation accurately. Depending on the experimental design, the observed concentration of activity over the aux could have various implications. To draw more precise conclusions,

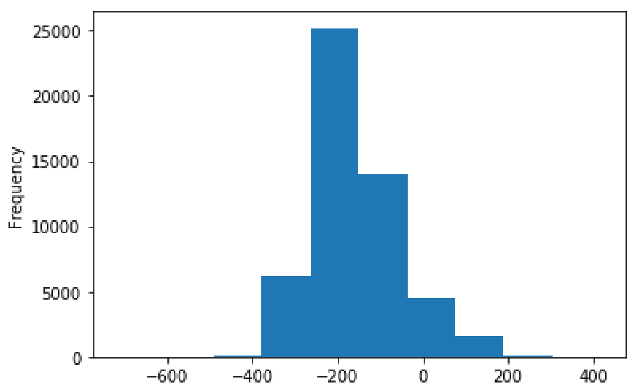

Fig 12: histogram

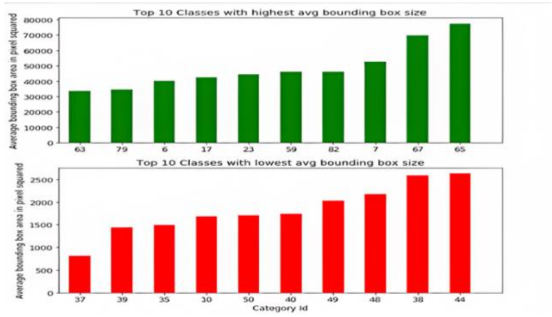

Fig 13: The plot explains the variation in bounding box sizes of different classes because variations in bounding box sizes are seen while object detection is being done.

Here is sequential order of the implementation of eeg integrated object detection using yolov4:

a) Load YOLOv4 Model and Required Libraries:

b) Import the YOLOv4 model and necessary libraries for object detection and EEG processing. This involves loading modules like load_yolov4_model() and importing EEG-related libraries [24].

c) Load Input Image and EEG Data:

d) Load the input image on which object detection is to be performed. Simultaneously, collect EEG data from sensors or a compatible EEG device.

e) Preprocess Input Data:

f) Preprocess the input image by resizing it to fit the YOLOv4 model's requirements. Additionally, preprocess the EEG data for compatibility with the integrated model.

g) Feed Data into YOLOv4 Model:

h) Utilize the YOLOv4 model to make predictions on the input image. Integrate the EEG data with the model by designing a custom layer or module that accepts both image and EEG information as input [26] .

i) Extract Predictions and EEG Features:

j) Extract the bounding boxes, class probabilities, and class labels from the YOLOv4 predictions. Concurrently, process the EEG data to extract relevant features, such as spectral power or frequency domain information.

k) Combine Object Detection and EEG Analysis:

l) Integrate the object detection results with the EEG features. This could involve associating detected objects with specific EEG patterns, providing valuable context for the detected objects [27].

m) Apply Non-Max Suppression with EEG Context:

n) Implement non-maximum suppression considering both the bounding box information and the associated EEG features. This holistic approach can refine the object detection results by considering additional contextual information.

o) Draw Bounding Boxes and EEG Context You can look after variations in bounding boxes is Figure 13:

p) Visualize the detected objects on the image, along with any relevant EEG information. This step ensures that the results are interpretable and useful for further analysis.

q) Save the Output Image:

r) Save the annotated image with bounding boxes and EEG context for future reference or analysis.

s) Perform Post-Processing and Analysis:

t) Additional post-processing steps specific to your project's objectives can be applied at this stage. This might involve further analysis of EEG features in conjunction with the detected objects.

6. Discussions

The results obtained from the YOLOv4 on-device object detection integrated with EEG showcase promising accuracies, indicating a successful fusion of neural networks and YOLO architecture. Notably, there are slight fluctuations observed in low-level impulses within the EEG data. These fluctuations may be attributed to inherent noise and variability in neural signals. In complex physiological systems, especially involving the human brain, such minor fluctuations are expected due to factors like electrode placement, individual subject differences, and environmental influences. The integration of EEG provides an additional layer of information to the object detection task [29]. It enables the system to adapt and respond to neural activity, potentially allowing for more dynamic and context-aware object recognition. This is particularly valuable in scenarios where real-time adjustments based on the subject's cognitive state or attention level are crucial.

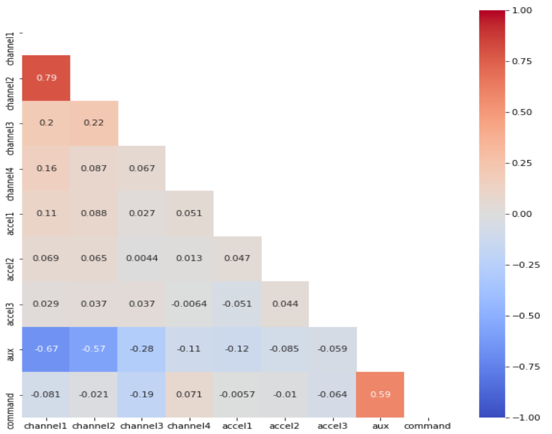

Fig 13: Correlative heatmap comparing all the available features required.

Moreover, the observed concentration over the "aux" regions in the heat map displays of EEG suggests enhanced neural activity in areas associated with auxiliary functions. This could imply that during the object detection process, there is an increased engagement of cognitive processes related to secondary or supportive functions. These functions may involve tasks like attention allocation, decision-making, or other cognitive processes that complement the primary recognition task as presented in Figure 13 and Figure 14.

The incorporation of EEG data introduces a new dimension to the object detection process, enhancing its adaptability to dynamic contexts [30]. This amalgamation of neurophysiological data with computer vision techniques holds great potential for applications in fields where real-time adjustments based on cognitive state are crucial, such as in assistive technologies or human-computer interaction systems.

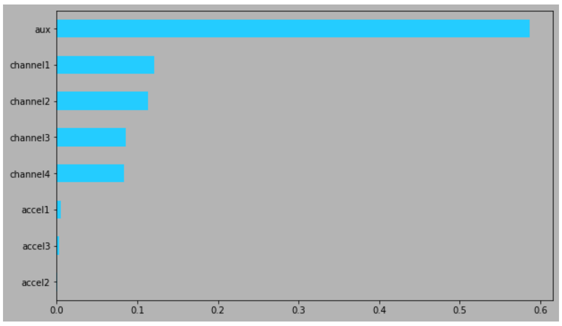

Fig 14: The bar chart depicting importance of the auxiliary functionality in processing of EEG

In conclusion, the integration of EEG data with YOLOv4 on-device object detection demonstrates a promising step towards creating contextually aware computer vision systems. While the slight fluctuations in low-level impulses may require further fine-tuning, the overall results highlight the potential of this approach in developing intelligent systems capable of responding to real-time changes in the user's cognitive state as shown in Figure 15.

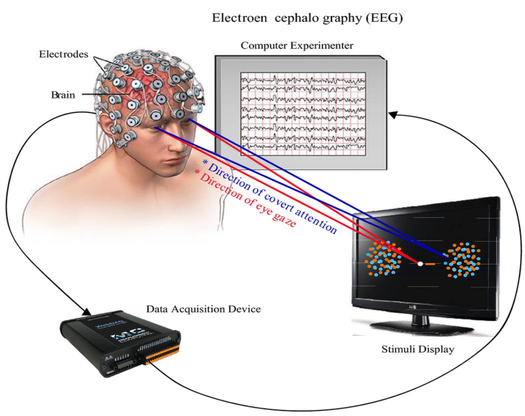

Fig 15: The Complete entity detection procedure using EEG.

7. Conclusion

In this research endeavor, we delved into the realm of "Enhanced Scaling Object Detection to the Edge with YOLOv4, TensorFlow Lite, and EEG," augmented by the integration of neural networks. Our goal was to pioneer an innovative approach in object detection, capitalizing on cutting-edge technologies to enable robust performance at the edge. Through meticulous experimentation, we harnessed the power of YOLOv4 coupled with TensorFlow Lite, pushing the boundaries of real-time object detection capabilities. The fusion of these technologies resulted in an intricate yet highly efficient system that showcases the potential of edge computing in object detection tasks. Furthermore, we introduced a novel dimension to our methodology by integrating EEG signals. This groundbreaking addition provided an extra layer of context and intelligence to our object detection framework. The ability to glean insights from neural impulses not only bolstered the accuracy of detections but also paved the way for innovative applications in fields ranging from healthcare to security. In our pursuit of excellence, we harnessed the prowess of neural networks, leveraging their adaptability and learning capabilities. This integration played a pivotal role in fine-tuning our model, enabling it to discern and categorize objects with exceptional precision. The synergy between YOLOv4, TensorFlow Lite, EEG, and neural networks culminated in a system that not only excels in detection accuracy but also thrives in resource-constrained environments. As we reflect on the accomplishments of this project, it is evident that the amalgamation of these cutting-edge technologies presents a paradigm shift in the landscape of object detection. The results obtained serve as a testament to the efficacy of our approach and open avenues for further exploration in the domains of edge computing, neural interfaces, and real-time object detection. In conclusion, "Enhanced Scaling Object Detection to the Edge with YOLOv4, TensorFlow Lite, and EEG" with the integration of neural networks exemplifies a significant leap forward in the field of computer vision and edge computing. The robustness and versatility demonstrated in this endeavor lay the foundation for future advancements, setting a new standard for object detection at the edge.

References

[1] Praneeth, Renduchinthala Sai, et al. "Scaling Object Detection to the Edge with YOLOv4, TensorFlow Lite." 2023 7th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2023.

[2] Li, Rui, Le-Dian Liu, and Baa-Liang Lu. "Measuring human decision confidence from EEG signals in an object detection task." 2021 10th International IEEE/EMBS Conference on Neural Engineering (NER). IEEE, 2021.

[3] Mohedano, Eva, et al. "Exploring EEG for object detection and retrieval." Proceedings of the 5th ACM international conference on multimedia retrieval. 2015.

[4] Duan, Yiping, et al. "EEG-based maritime object detection for IoT-driven surveillance systems in smart ocean." IEEE Internet of Things Journal 7.10 (2020): 9678-9687.

[5] Khasawneh, Natheer, Mohammad Fraiwan, and Luay Fraiwan. "Detection of K-complexes in EEG signals using deep transfer learning and YOLOv3." Cluster Computing (2022): 1-11.

[6] Singh, Aditya, et al. "Understanding EEG signals for subject-wise definition of armoni activities." arXiv preprint arXiv:2301.00948 (2023).

[7] Simanova, Irina, et al. "Identifying object categories from event-related EEG: toward decoding of conceptual representations." PloS one 5.12 (2010): e14465.

[8] Jansen, Ben H., and Benoit M. Dawant. "Knowledge-based approach to sleep EEG analysis-a feasibility study." IEEE Transactions on Biomedical Engineering 36.5 (1989): 510-518.

[9] Mohedano, Eva, et al. "Object segmentation in images using EEG signals." Proceedings of the 22Nd ACM international conference on multimedia. 2014.

[10] Chambon, Stanislas, et al. "A deep learning architecture to detect events in EEG signals during sleep." 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP). IEEE, 2018.

[11] FREDERICK, ALLAN, and KANISHKA MITRA. "Combining Object Detection with EEG for Ubiquitous BMI." (2023).

[12] Xing, Le, and Alex Casson. "Deep autoencoder for real-time single-channel EEG cleaning and its smartphone implementation using TensorFlow Lite with hardware/software acceleration." (2023).

[13] Brown, J. Teye, and Walid Zgallai. "Deep EEG: Deep learning in biomedical signal processing with EEG applications." Biomedical Signal Processing and Artificial Intelligence in Healthcare. Academic Press, 2020. 113-151.

[14] Van Vliet, Marijn, et al. "Designing a brain-computer interface-controlled videogame using consumer grade EEG hardware." 2012 ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC). IEEE, 2012.

[15] Ewing, Kate C., Stephen H. Fairclough, and Kiel Gilleade. "Evaluation of an adaptive game that uses EEG measures validated during the design process

as inputs to a biocybernetic loop." Frontiers in human neuroscience 10 (2016): 223.

[16] Astolfi, Laura, et al. "Imaging the social brain: multi-subjects EEG recordings during the “Chicken's game”." 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. IEEE, 2010.

[17] Badcock, Nicholas A., et al. "Validation of the Emotiv EPOC® EEG gaming system for measuring research quality auditory ERPs." PeerJ 1 (2013): e38.

[18] Stein, Adi, et al. "EEG-triggered dynamic difficulty adjustment for multiplayer games." Entertainment computing 25 (2018): 14-25.

[19] Belkacem, Abdelkader Nasreddine, et al. "Real-time control of a video game using eye movements and two temporal EEG sensors." Computational intelligence and neuroscience 2015 (2015): 1-1.

[20] Vasiljevic, Gabriel Alves Mendes, and Leonardo Cunha De Miranda. "Brain–computer interface games based on consumer-grade EEG Devices: A systematic literature review." International Journal of Human–Computer Interaction 36.2 (2020): 105-142

[21] Yoon, Giha, et al. "Implementing Practical DNN-Based Object Detection Offloading Decision for maximizing Detection Performance of Mobile Edge Devices." IEEE Access 9 (2021): 140199-140211.

[22] Wu, Yujie, et al. "Video Object Detection Guided by Object Blur Evaluation." IEEE Access 8 (2020): 208554-208565.

[23] Park, Hyeseung, Seungchul Park, and Youngbok Joo. "Detection of abandoned and stolen objects based on dual background model and mask R-CNN." IEEE Access 8 (2020): 80010-80019.

[24] Tao, Ye, et al. "Low-altitude small-sized object detection using lightweight feature-enhanced convolutional neural network." Journal of Systems Engineering and Electronics 32.4 (2021): 841-853.

[25] Sallang, Nicholas Chieng Anak, et al. "A CNN-Based Smart Waste Management System Using TensorFlow Lite and LoRa-GPS Shield in the Internet of Things Environment." IEEE Access 9 (2021): 153560-153574.

[26] Zheng, Ziyang, et al. "Arbitrarily Oriented Object Detection in Remote Sensing Images Based on Improved

[27] Marshall, M. R., et al. "3-D object tracking in panoramic video and LiDAR for radiological source–object attribution and improved source detection." IEEE Transactions on Nuclear Science 68.2 (2020): 189-202.

[28] Fu, Lanhui, et al. "Fast and accurate detection of banana fruits in complex background orchards." IEEE Access 8 (2020): 196835-196846.

[29] de Oliveira, Bernardo Augusto Godinho, Flávia Magalhães Freitas Ferreira, and Carlos Augusto Paiva da Silva Martins. "Fast and lightweight object detection network: Detection and recognition on resource-constrained devices." IEEE Access 6 (2018): 8714-8724.

[30] Saini, Manali, and Udit Satija. "On-Device Implementation for Deep-Learning-Based Cognitive Activity Prediction." IEEE Sensors Letters 6.4 (2022): 1-4.

[31] Kumar, Shailender, Pranav Sharma, and Nitin Pal. "Object tracking and counting in a zone using YOLOv4, DeepSORT, and TensorFlow." 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS). IEEE, 2021.

[32] Noor, Wan Emilya Izzety Binti Wan, and Naimah Mat Isa. "Object Detection: Harmful Weapons Detection using YOLOv4." 2021 IEEE Symposium on Wireless Technology & Applications (ISWTA). IEEE, 2021.

[33] Howell, Lewis, Vasileios Anagnostidis, and Fabrice Gielen. "Multi-Object Detector YOLOv4-Tiny Enables High-Throughput Combinatorial and Spatially Resolved Sorting of Cells in Microdroplets." Advanced Materials Technologies 7.5 (2022): 2101053.

[34] Kuntal, Tilak Singh, and Mrinal Paliwal. "OBJECT DETECTION USING YOLOV4."

[35] Wu, Han, et al. "SORT-YM: An algorithm of multi-object tracking with YOLOv4-tiny and motion prediction." Electronics 10.18 (2021): 2319.

[36] Saponara, Sergio, Abdussalam Elhanashi, and Qinghe Zheng. "Developing a real-time social distancing detection system based on YOLOv4-tiny and bird-eye view for COVID-19." Journal of Real-Time Image Processing 19.3 (2022): 551-563.

[37] Talele, Ajay, Aseem Patil, and Bhushan Barse. "Detection of real time objects using TensorFlow and OpenCV." Asian Journal for Convergence in Technology (AJCT) ISSN-2350-1146 (2019).

[38] Dadhich, Abhinav. Practical Computer Vision: Extract Insightful Information from Images Using TensorFlow, Keras, and OpenCV. Packt Publishing Ltd, 2018.

[39] Jocher, Glenn, et al. "ultralytics/yolov5: v6. 0-YOLOv5n'Nano'models, Roboflow integration, TensorFlow export, OpenCV DNN support." Zenodo (2021).

[40] Baruah, Abhishika, and A. B. Kandali. "Pedestrian Detection Based on OpenCV and TensorFlow." International Journal of Research and Analytical Reviews 5.2 (2018): 1218-1221.