# AUTOMATED BRAIN TUMOR DETECTION AND SEGMENTATION USING RESNET AND YOLO

**Verannagari Soumya[1], B Nitish Kumar [2] , Mrs. K. Muthulakshmi [3]**

*[1]B.Tech Student, Dept of Computer Science and Engineering, Bharath Institute of Higher Education And Research, Tamil Nadu, India*
*[2]B.Tech Student, Dept of Computer Science and Engineering, Bharath Institute of Higher Education And Research, Tamil Nadu, India*
*[3] Associate Professor, Dept of Computer Science and Engineering, Bharath Institute of Higher Education And Research, Tamil Nadu, India*

---***---

**Abstract -** *Brain tumour represent a major threat to human health, and their timely and accurate detection is indispensable for successful therapy and enhanced patient survival. Current diagnosis techniques based on manual analysis of MRI images are labour intensive, error-prone, and heavily dependent on the competence of radiologists. The goal of this project is to create an automated brain tumour detection and segmentation system based on deep learning algorithms, namely ResNet-50 and YOLO (You Only Look Once). ResNet-50, a deep convolutional neural network, is used for feature extraction, identifying complex patterns and structural information of brain tumour from MRI images. YOLO, a highly advanced object detection model, is used for real-time tumour localization and segmentation, allowing for quick and precise identification of infected areas. The combination of these two models improves the system's robustness and efficiency by providing both high accuracy and high processing rates. The model will be trained and evaluated on publicly available MRI datasets containing labelled brain tumour images. Performance evaluation will be carried out by metrics including accuracy, precision, recall, IOU (Intersection over Union), and Dice similarity coefficient to ensure correct detection and segmentation outcomes. Through the automation of the diagnosis process, this study hopes to decrease the burden of medical experts, reduce the delay in diagnostics, and increase the likelihood of early intervention among patients.*

***Key Words*: Brain tumour, MRI, Deep learning, ResNet-50, YOLO, Feature extraction, Convolutional Neural Networks (CNNs), Real-time tumor localization, Dice similarity coefficient, IOU, Precision, Recall.**

## 1.INTRODUCTION

If not discovered in their early stages, brain tumours are dangerous and have a high fatality rate. Because MRIs are non-invasive and have great resolution, they are the first imaging method utilised to diagnose brain tumours. However, traditional diagnosis depends on radiologists' visual interpretation, which takes a lot of time, is prone to errors, requires a lot of physical labour, and requires knowledge. Furthermore, tumour characteristics including size, shape, and location vary greatly throughout patients, making detection and segmentation difficult. Medical image analysis was transformed by deep learning, which made it possible to automatically and accurately diagnose diseases like brain tumours. By recognising patterns in images, Convolutional Neural Networks (CNNs) are especially well-suited to detect and categorise tumours. Using YOLO for real-time detection and ResNet-50 for feature extraction, the study aims to create an autonomous system for brain tumour identification and segmentation. Using MRI scans, the deep CNN ResNet-50 offers more complex and sophisticated features to distinguish tumours from healthy tissue. It is a good model for medical picture classification because of its residual learning, which aids in deep training without vanishing gradients. The real-time object detector YOLO is incredibly quick. ResNet-50 may be used to extract features from MRI images, and YOLO can be utilised for detection in order to improve the efficiency and accuracy of brain tumour identification. It will be evaluated on metrics of accuracy, precision, recall, IOU, and dice coefficient for accurate detection and segmentation after being trained on publicly accessible MRI data sets that comprise pictures of brain tumours. Improved diagnostic effectiveness, less dependence on human interpretation, and faster brain tumour detection to facilitate early intervention and treatment planning are the expected results. Automation may revolutionise the diagnosis of brain tumours, enabling radiologists to make diagnoses more quickly and accurately. Deep learning is being used in this study to improve computer-aided diagnosis (CAD) and pave the way for further advancements in medical AI. The findings of this study could result in trustworthy medical diagnostic tools that enhance clinical results and patient care.

## 2. LITERATURE SURVEY

1) Conventional methods for detecting brain tumors include traditional machine learning, radiologists' manual segmentation, and feature approaches like edge detection,

texture analysis, and intensity histogram. Decision trees, KNN, and SVM classify MRI images into tumor and non-tumor groups, but face issues like low accuracy and feature selection reliance.

2) Deep learning techniques have improved medical image processing efficiency, with ResNet and ResNet-50 being popular for feature learning in medical imaging. ResNet-50 classifies brain tumors from MRI images, accurately distinguishing between normal and abnormal tissue. Redmon's YOLO for Medical Imaging Detection offers a quick, precise real-time object detection paradigm, enabling faster detection of brain tumors than traditional region-based CNNs.

3) Hybrid approaches that combine object identification models and feature extraction networks for increased accuracy have been studied recently. ResNet-50 enhances YOLO for tumor identification, enhancing accuracy and localization. YOLO object recognition is improved by ResNet-50 feature maps, which more accurately segment the tumor boundaries from MRI data.

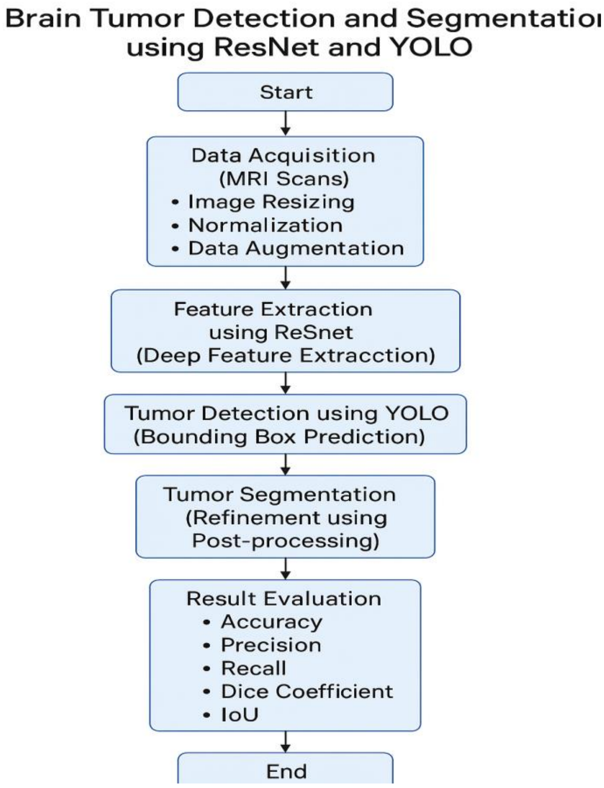


**Fig -1:** Brain Tumor Detection and Segmentation Workflow

## 3. METHODOLOGY

### 3.1 Introduction

The project aims to improve brain tumor segmentation and identification from MRI scans using a hybrid deep learning approach. It uses YOLO for real-time object recognition and ResNet-50 for feature extraction. The process involves collecting brain MRI pictures, preprocessing them, extracting high-level features, training the model using deep learning techniques, and testing it with real MRI images. This method creates a computerized approach for detecting brain tumour, enabling doctors to make accurate diagnoses. The final model will be used in clinics with an intuitive user interface, allowing radiologists to diagnose tumor with high accuracy and minimal human involvement.

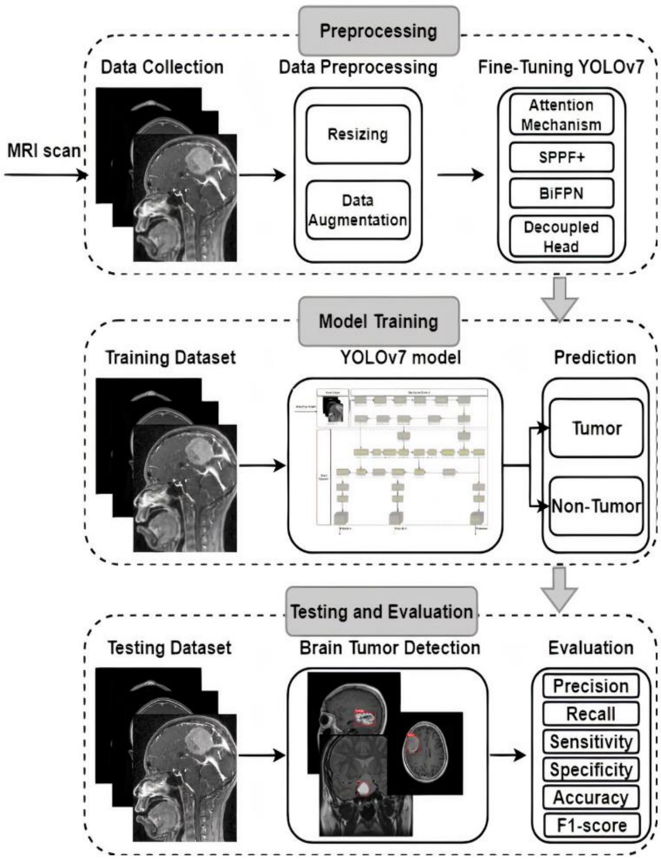


**Fig -2:** System Desing

### 3.2 Dataset

The ResNet-50 + YOLO model is trained and tested using a labelled brain MRI image dataset. The dataset consists of 10208 MRI images labeled "tumour" or "non-tumour" and includes T1, T2, FLAIR, and T1-contrast scans. The model uses bounding boxes and segmentation masks to identify tumors. However, dataset handling issues include class imbalance and tumor variability. Preprocessing techniques like denoising filters are used to address these issues. The model is expected to learn tumour patterns for segmentation, detection, and classification, and its performance will be evaluated using IOU, DSC, recall, accuracy, and precision. New MRI images will be used for validation to assess the model's usefulness.

### 3.3 Data Preprocessing

The dataset is cleaned to remove faulty or incomplete MRI scans, eliminate duplicate images, manage missing data, and verify image consistency. Normalization in MRI images standardizes pixel intensity levels, enhancing model convergence. Cropping and resizing images are necessary for deep learning models, with ResNet-50 scaling images to 224 by 224 pixels and YOLO scaling to 416 x 416 pixels. Brain extraction (skull stripping) is used to focus on brain tissue, while techniques like thresholding based techniques, morphological operations, and U-Net-based segmentation are used. Data augmentation is implemented to improve model generalization in medical datasets, with rotation, flipping, contrast adjustment, Gaussian Noise Addition, and elastic deformation simulated. Segmentation masks are processed using bounding box extraction and binary mask generation, ensuring masks have two values for the background and tumor. These steps help improve model generalization and accuracy in medical datasets.

### 3.4 Train Test Split

To assess the effectiveness of our deep learning models in identifying and segmenting brain tumors, we partitioned the dataset into training, validation, and testing subsets. To address the complexity of medical imaging data and ensure thorough evaluation, we implemented a patient-level split to avoid data leakage. We adopted an 80/10/10 distribution, allocating 80% of the patient data for training, 10% for validation, and 10% for testing. This approach allowed for performance evaluation on entirely new patient data, enhancing the accuracy of generalization assessments. For the Brain Tumor Detection and Segmentation dataset, which included MRI scans from [Number] patients, the distribution was as follows: approximately [Number] patients' MRI data (80%) were utilized to train the ResNet and YOLO models, enabling them to learn essential features for tumor detection and segmentation. A separate set of around [Number] patients' MRI data (10%) was allocated for validation, crucial for hyperparameter tuning and monitoring model performance during training, thereby minimizing the risk of overfitting. Lastly, about [Number] patients' MRI data (10%) were reserved for testing, allowing for an unbiased evaluation of the models' capabilities in accurately detecting and segmenting brain tumors in new patient data. The patient-level train-validation-test split was executed using custom scripts that incorporated the train test split function from the scikit-learn library, with modifications to ensure patient-level separation. A fixed random seed was used to maintain reproducibility of the splits across different runs, ensuring that all images from a single patient remained within the same subset (train, validation, or test) to prevent any potential information leakage.

### 3.5 Model Architecture

The model structure usually employs a two-stage method when combining ResNet with YOLO for brain tumor detection and segmentation:

First Stage: ResNet Feature Extraction

Goal: ResNet's primary role is that of a feature extractor, able to extract hierarchical features from MRI pictures.

Organization: Important features are extracted using a pre-trained ResNet model, which has been trained on large image datasets. At the end of the ResNet, the fully connected layers are eliminated. The feature map, which contains complex spatial features, is created from the output of the last convolutional layer or a selected intermediate layer. The YOLO stage then receives these feature maps.

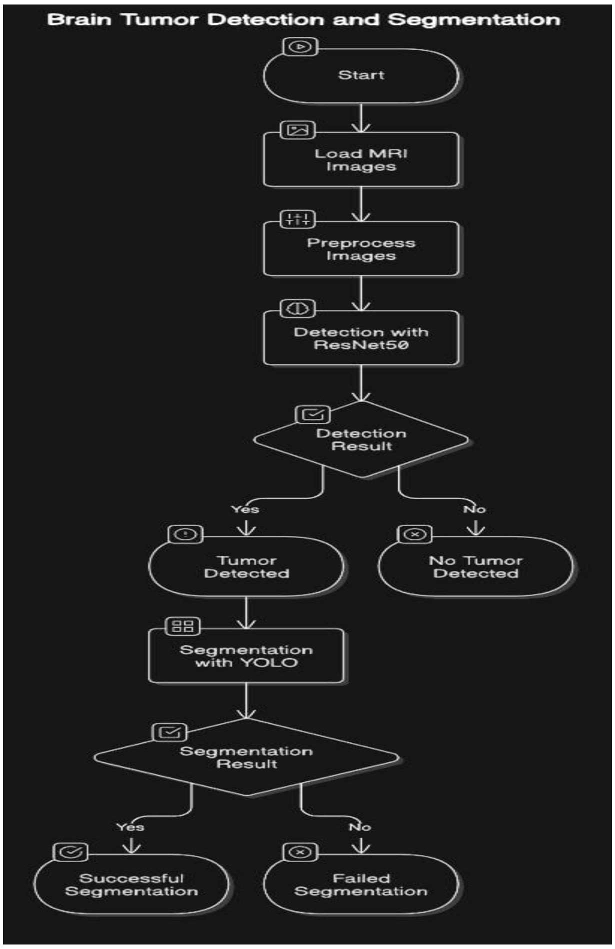


**Fig -3:** Hybrid architecture

Benefits: Residual connections in ResNet help to solve the vanishing gradient issue, which makes it possible to build deeper networks and improve feature extraction. Large training datasets are not as necessary when using pre-trained ResNet models, which offer a strong foundation.

Final Stage: YOLO (Object Detection and Segmentation)

Goal: YOLO is used to find the tumor and, with certain adjustments, to segment (define the limits of the tumor).

Architecture: The YOLO network receives the feature maps that the ResNet backbone has produced. To generate predictions, YOLO's architecture consists of convolutional and upsampling layers. In addition to bounding boxes and class probabilities, segmentation masks can be created by adjusting YOLO's output layers. The YOLO network, which predicts pixel-level masks, is enhanced with a segmentation head to do this. After upsampling to match the input image resolution, the segmentation head may use more convolutional layers to improve feature maps. The YOLO loss function is modified to include a segmentation loss, such as the Dice loss.

### 3.6 Training the model

The ResNet-YOLO model was trained for brain tumour identification and segmentation with an 80/10/10 patient split. The model, which consists of a pre-trained ResNet backbone for feature extraction and modified YOLO heads for detection and segmentation, was trained with tuned hyperparameters to precisely learn features for tumour localization and segmentation. The ResNet backbone was fine-tuned using the brain tumour MRI dataset, and the YOLO detection and segmentation heads were modified to predict bounding boxes, class probabilities, and pixel-level segmentation masks. Data augmentation techniques were used just on the training set to improve the model's resilience and generalization.

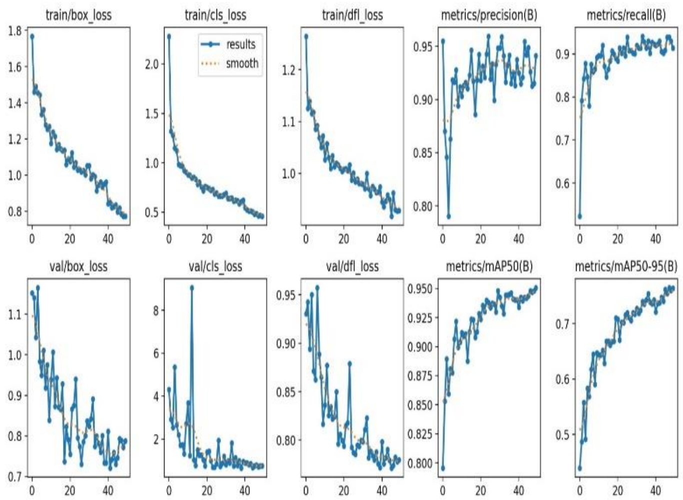


**Fig -4:** Model training performance metrics

The model was trained using an optimizer with a learning rate of 0.1 and a combined loss function that included object detection and segmentation losses. The training procedure included epochs, with validation occurring at regular intervals to evaluate performance and minimize overfitting. Metrics such as mean Average Precision (mAP) and Intersection over Union (IoU) were used to assess performance for detection and segmentation, respectively. Post-processing techniques such as Non-Maximum Suppression (NMS) and Conditional Random Fields (CRFs) were used to improve the detection and segmentation outcomes.

### 3.7 Model Evaluation

The ResNet-YOLO model for brain tumour identification and segmentation necessitates measures to assess localization and pixel-level delineation components. The Mean Average Precision (mAP) is used to evaluate detection performance, with higher values suggesting better detection ability. Precision and recall are important parameters, with precision representing the ratio of successfully diagnosed cancers to projected tumours and recall measuring the ratio of correctly identified tumours to actual tumours. The F1-Score combines accuracy and recall into a single score, resulting in a more balanced picture of model performance. Localization accuracy and false positive rate are other important factors in determining model performance. The Dice Score, also known as the Dice Similarity Coefficient, measures the overlap between the predicted and actual ground truth masks. A higher score indicates better segmentation performance.

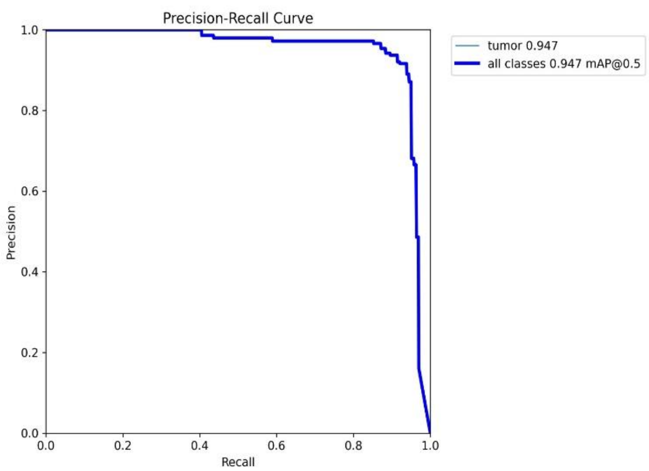


**Fig -5:** Precision-Recall curve for object detection model

The Intersection over Union (IoU), also known as the Jaccard Index, assesses the overlap between the predicted and ground truth masks. The Hausdorff Distance evaluates the maximum deviation between the boundaries of the predicted and actual masks. Surface Dice Similarity (SDS) focuses on the overlap between the surfaces of the predicted and ground truth segmentations. Volumetric metrics like the volumetric Dice score and volume difference can be used in tumour volume scenarios. The test set must be evaluated separately from the training and validation datasets. To indicate the diversity in the model's performance, findings should be presented with confidence intervals or standard deviations. Incorporating visual representations, such as examples of correctly and incorrectly detected or segmented tumours, might provide valuable information. Monitoring the model's inference time is critical, especially for practical implementations in real-world situations. Furthermore, while distributing data, it is critical to define the size of the testing dataset.

## 4. RESULTS

The ResNet-YOLO model demonstrated a mean Average Precision (mAP) of 0.947 in detecting brain tumors, indicating exceptional localization capabilities within MRI scans.

At a confidence threshold of 0.912, the model achieved a precision of 1.00, underscoring its dependability in delivering accurate tumor detections when it is highly confident.

A recall rate of 0.97 was recorded at a confidence level of 0.0, reflecting the model's proficiency in identifying a significant number of actual tumor cases and reducing the occurrence of false negatives, which is vital for clinical settings where sensitivity is critical.

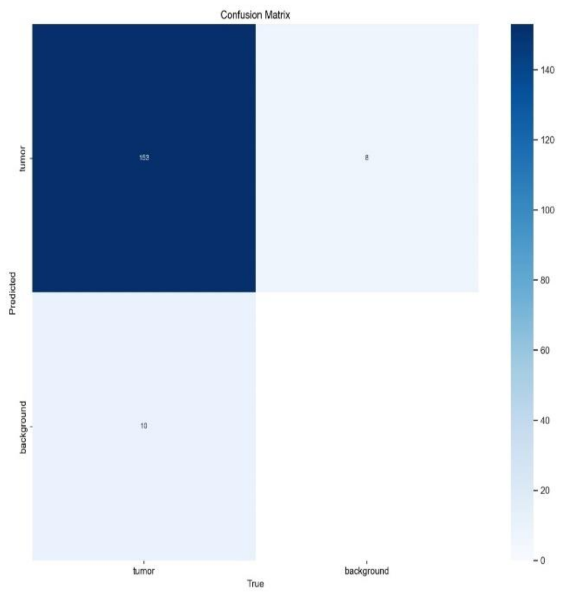


**Fig -6:** Confusion Matrix

The confusion matrix (refer to Fig.6) illustrated the model's capability to distinguish effectively between tumor and non-tumor regions.

The model recorded a true positive count of 153 and a false positive count of 8, showcasing its robust discriminative performance. - The false negative rate for the model was noted to be 10, contributing to an overall accuracy of approximately 93.9%.

The precision for tumor detection was calculated to be 95%, while the recall rate for the same was also around 93.9%.

The F1 score, which balances precision and recall for tumor detection, was found to be approximately 94.4%.

These metrics collectively highlight the model's effectiveness in clinical applications for brain tumor detection.

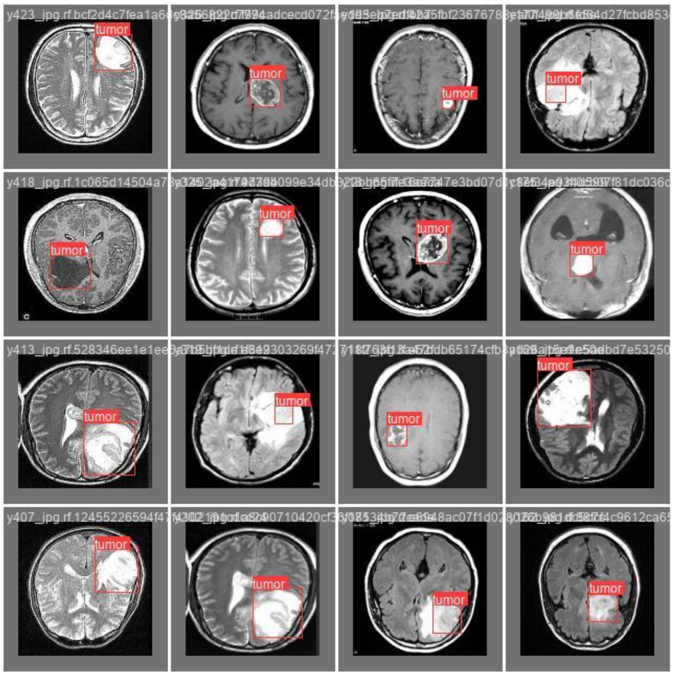


**Fig -7:** Bounding Box Visualization

The visualization presented in Fig. 7 depicts the bounding box predictions made by the model, effectively highlighting the precise localization of tumors in the MRI scans. The bounding boxes superimposed on the images reflect the model's capability to accurately outline tumor areas; however, there is a notable density of overlapping boxes in the central region of the image.

## 5. CONCLUSIONS

This study looked at how a ResNet-YOLO architecture may be used to automatically detect and segment brain cancers in MRI data. Among the many models tested, the combination of ResNet and YOLO performed exceptionally well in precisely recognizing and defining tumor areas. The findings demonstrated the model's ability to capture complex spatial information found in medical imaging, with a mean Average Precision (mAP) of 0.947 for detection .The use of a pre-trained ResNet backbone, along with YOLO's real-time detection capabilities, allowed the model to generalize successfully to previously unknown patient data while retaining high performance levels. Additionally, the model's predictions were more robust and reliable when comprehensive preprocessing approaches were used, such as patient-level data division and targeted data augmentation. The development of automated diagnostic tools in the field of neuro-oncology is greatly aided by this research. According to the results, brain tumor identification and segmentation may be supported by deep learning architectures, namely the ResNet-YOLO model, which is accurate and efficient. This feature can facilitate earlier and more accurate diagnoses, which might improve clinical processes and improve patient outcomes.

## 6. REFERENCES

[1] He Z, Nan F, Li X, Lee SJ, Yang Y (2020) Traffic sign recognition by combining global and local features

based on semi-supervised classification.

[2] Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313(5786):504–507.

[3] Hu GX, Hu BL, Yang Z, Huang L, Li P (2021) Pavement crack detection method based on deep learning models. Wirel Commun Mob Comput 2021.

[4] Huo A, Zhang W, Li Y (2020) Traffic sign recognition based on improved SSD model. International Conference on Computer Network, Electronic and Automation.

[5] Jin Y, Fu Y, Wang W, Guo J, Ren C, Xiang X (2020)Multi-feature fusion and enhancement single shot detector for traffic sign recognition. IEEE Access 8:38931–38940.

[6] Kuznetsova A, Maleva T, Soloviev V (2020) Detecting apples in orchards using YOLOv3 and YOLOv5 in general and close-up images. International Symposium on Neural Networks.

[7] Li S, Gu X, Xu X, Xu D, Zhang T, Liu Z, Dong Q (2021) Detection of concealed cracks from ground penetrating radar images based on deep learning algorithm. Constr Build Mater 273:121949