Feature Extraction for Image Analysis and Detection using Machine Learning Techniques

Article · January 2023

DOI: 10.35444/IJANA.2023.14401

2 authors, including:

Muath Sabha
Arab American University
24 PUBLICATIONS **120** CITATIONS

SEE PROFILE

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

Feature Extraction for Image Analysis and Detection using Machine Learning Techniques

Adel Hassan
Faculty of Engineering and Information Technology, Arab American University, Jenin, Palestine
Email: adel7377@gmail.com
Muath Sabha
Faculty of Engineering and Information Technology, Arab American University, Jenin, Palestine
Email: muath.sabha@aaup.edu

----------**ABSTRACT**----------
Feature extraction is the most vital step in image classification to produce high-quality and good content images for further analysis, image detection, segmentation, and object recognition. Using machine learning algorithms, profound learning like convolutional neural network CNN became necessary to train, classify, and recognize images and objects like humans. Combined feature extraction and machine learning classification to locate and identify objects on images can then be an input of automatic recognition systems ATR such as surveillance systems CCTV, to enhance these systems and reduce time and effort for object detection and recognition in images based on digital image processing techniques especially image segmentation that differentiate from computer vision approach. This article will use machine learning and deep learning algorithms to facilitate and achieve the study's objectives.

Keywords - **Automatic Recognition Systems, Convolution Neural Network, Digital Image Processing, Feature Extraction, Image Segmentation.**

Date of Submission: 1/11/2022 Date of Acceptance: 1/12/2022

I. INTRODUCTION

This article aims to produce high-quality images as input for an automatic target recognition system (ATRS) for multiple industries, such as surveillance systems CCTV, unmanned Flight systems, and others [1].

Feature extraction is the crucial image classification step to producing high-quality and good content images. The image processing flow consists of three phases. The First phase is feature extraction [2]; we will use different methods such as histogram gradients, HOG [3], and convolution neural network CNN [4].

The second phase is the classification mode used as support vector machine SVM [5] and Random Forest RF ensemble learning methods [6].

In the third phase, we will use clustering methods such as structural similarity index matrix SSIM [7] to put output images into similar groups to begin the final step, which is to analyze and recognize the nature of objects in the images to be as input for automated recognition system used in applications such mentioned above for CCTV system [27] and other systems that required to detect and recognize the objects on images.

II. BACKGROUND

The essential Image processing key stages are object recognition/detection and image segmentation.

For differences between object detection and image segmentation, the process begins with object recognition that is split into two operations: image localization and image classification. Then we can merge the two last operations to form the object detection stage; after the object has been identified in the image, we use image segmentation to recognize what the object is. As shown in Figure-1.

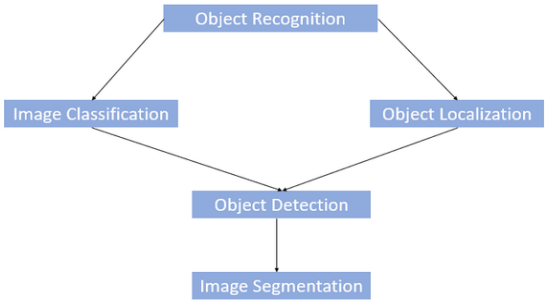

Fig 1. Object Detection Vs. Image Segmentation

II.I OBJECT RECOGNITION OR

Object recognition is the way to identify the object in images, as shown in Fig 3. Then machine learning and ML techniques can then recognize which type of object is an illustration of how humans do if it is animal or human, and so on.

Fig 2. Object Recognition

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

II.I.I Object Recognition Using Machine Learning (ML)

1. HOG takes samples of images that contain object and do not contain object by using an image histogram and then train the ML model like SVM.
2. Bag of features model: represents an image as an orderless collection of image features like Scale-invariant feature transform SIFT [8].
3. Viola-Jones algorithm: Extract vast numbers of features in the image using Haar-like, then pass them to a classifier for the image detection process. It is used in most detecting faces on images near to real-time.

II.I.II Object Recognition Using Deep Learning (DL)

- In most cases, the convolution Neural Network (CNN) is used for object recognition and image classification. The object can be identified by comparing the input of the image and the output using multiple classes in terms probability of other classes in the model with higher values and dropping lower values. This way can do image feature extraction without comparing existing data as ML does, like R-CNN, Fast R-CNN, SSD, and Yolo v3.

II.I.III Challenges of Object Recognition (OR)

- Because CNN depends on the output of classes and consists of one layer for each class, the model will probably not work when there are different classes in the image. Moreover, localization issues arise since output classes need to be labeled for objects and the bounding box location of objects on the image.

II.II IMAGE Classification (IC)

- Image labeled as a class for the object using different metrics such as accuracy or probability of output classifier for input images. The dog has been labeled the class of dog with a high probability of up to 97%. As shown in Fig 3.

Fig 3. Image Classification

- Classification Algorithms such as SVM, RF, and deep learning algorithms Convolutional Neural Networks for pre-trained like Vgg 16, Inception, ResNet [9], Stochastic Depth ResNet, exception, Mobile Net ..etc.
- Four top State-of-the-Art pre-trained models for image classification, as shown in Fig. 4.

Model	Year	Number of Parameters	Top-1 Accuracy
VGG-16	2014	138 Million	74.5%
ResNet-50	2015	25 Million	77.15%
Inception V3	2015	24 Million	78.8%
EfficientNetB0	2019	5.3 Million	76.3%
EfficientNetB7	2019	66 Million	84.4%

Fig 4. Pre-trained models for image classification

II.III Object Localization (OL)

- This is done by locating the object's existence on the image using the dimensions of width and high of an object with a bounding box.

II.IV Object Detection (OD)

It uses object localization and image classification techniques to produce a bounding box of objects in the image with a class label. As shown in Fig. 5, the differences between these techniques detect and classify objects and images. But, there are some challenges using this approach when an object contains curves or does not unform shape because it deals with a bounding box for the object, so this makes it, in some cases, not accurate when there are multiple objects in the image that are very close to each other.

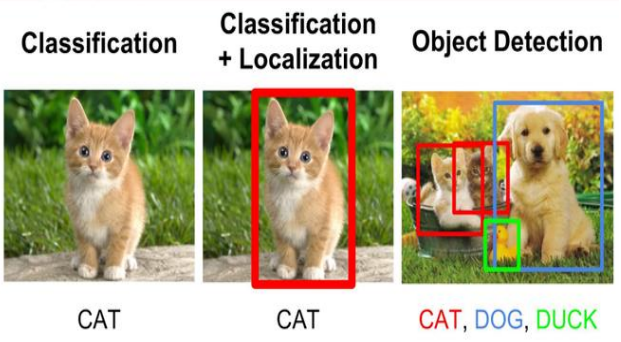

Fig 5. Object classification and detection boxes

II.V Image Segmentation (IS)

- This technique is similar to object detection without using a bounding box but using pixel-wise masks for the object on the image. So, it can help remove the challenges of object detection when objects have non-uniform shapes like curves. We will explain more in the next section using Mask R-CNN [10].
- Image segmentation consists of two methods, first is instance segmentation which colors the object based on different labeled pixels. The second is semantic segmentation, which colors the object-based category class labeled. As shown in fig. 6 shows the differences between object detection and image segmentation. And Fig. 7

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

shows differences between instant and semantic segmentation.

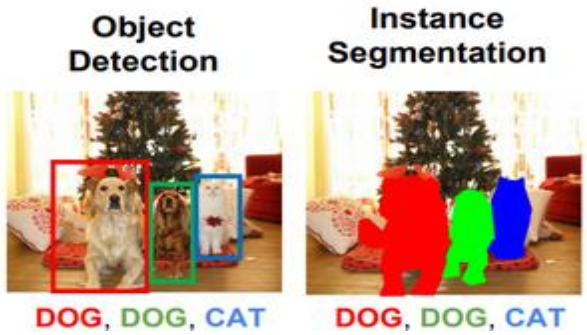

Fig 6. Object detection vs. Instant segmentation

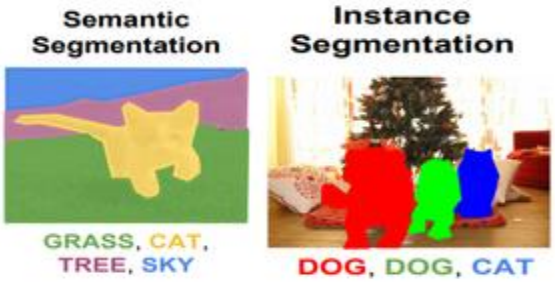

Fig 7. Semantic Segmentation vs. Instant Segmentation

- Using image segmentation to divide the images into small groups of segments can reduce the complicity by dealing with a simple image for more analysis, and that be done pixel-wise for each element on an image with labeled categories. Image segmentation has two image segmentation approaches and five image segmentation techniques described as follows:

II.V.I IMAGE SEGMENTATION APPROACHES

1. Similarity approach:
This approach is similar to the clustering approach by putting each pixel into a segment based on a specific threshold using the similarity between these segments.

2. Discontinuity approach:
Unlike similarity using a threshold for pixels, it uses pixel values intensity like point, line, and edge detection techniques.

II.V.II IMAGE SEGMENTATION TECHNIQUES

1. Threshold-Based Segmentation:
These techniques depend on pixel intensity threshold values to build binary images and use dynamic threshold techniques [11].
2. Edge-Based Segmentation EBS:
These techniques depend on edges operators like color or gray levels, So edges can be marked on images even when images are in a continuous mode based on discontinuity in color or gray levels.
3. Region-Based Segmentation RBS:
Divide regions into groups based on similar pixel properties such as color and intensity. This technique can work more when there are a lot of noises in images. In addition, it has two types, one called Region growing and the second called split & merge Region.
4. Clustering-Based Segmentation CBS:
The most popular methods of this technique are K-means, which divides each element into segments based on groups using k of clusters and the distance of each component from the center of each k.
5. Artificial Neural Network-Based Segmentation - Mask R-CNN.

- There is more than one technique using ANN-based segmentation, as follows:

I. Convolutional Neural Networks (CNN)
II. Region-Based Convolutional Neural Networks (R-CNN)
III. Faster R-CNN with Region Proposal Networks (RPN)
IV. Mask R-CNN

I. Neuron in one layer is connected to CNN is the most popular artificial neural network ANN technique used for images to optimize the processing of image pixel data called a convolutional neural network and consists of three main layers, as shown in Fig. 8; the first layer is the convolution layer use kernels and filters to produce feature amp of the input image. The second layer, called pooling, summarizes feature maps into patches to help create a feature map sample and forward it to the next layer. The third layer is the "fully" connected layer, ensuring every neuron in other layers can identify and recognize the object on the image.

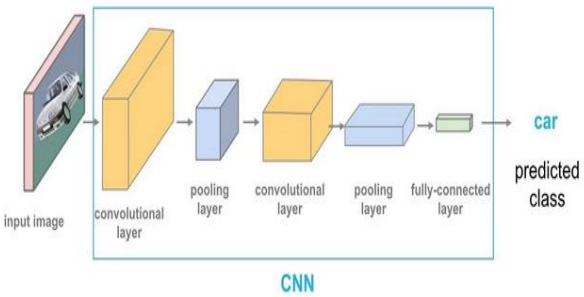

Fig 8. CNN architecture

II. R-CNN, a Region-based convulsion neural network [12], uses the regions of interest approach to identify the object on the image using bounding boxes and then put them into different labeled classes, as shown in Fig. 9.

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

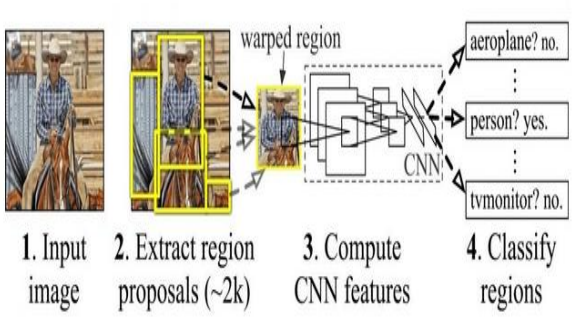

Fig 9. R-CNN architecture

III. Faster R-CNN is an improved region-based convolution neural network method with two stages [13]. The first one is called Region proposed network PRN, and the name comes from the fact that it proposes many objects that exist in an image, as shown in Fig.10.
- In the second stage, called Region of interest pooling, Rolpool extract features on images using bounding box regression for the object after the classification process; as shown in Fig. 11, that feature map is extracted using Ropool without the need to filter only the object within the Region of interest, and this is why it considerably faster to identify and recognize the object on image than R-CNN.

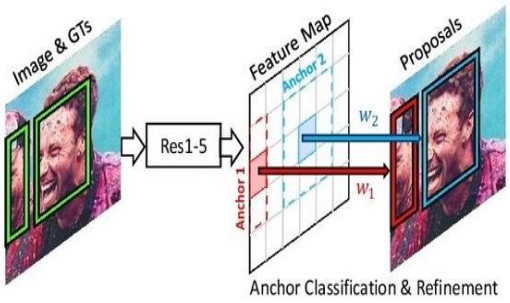

Fig 10. Region of interest pooling

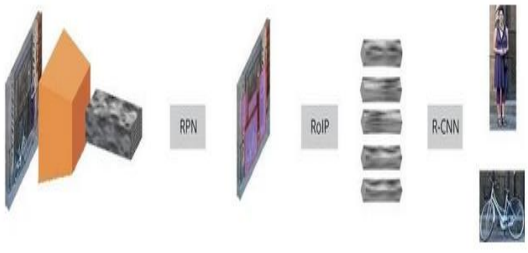

Fig 11. Fast CNN architecture

IV. Mask R-CNN consists of the most fit used for image segmentation since it locates and identifies objects and their boundaries like points, lines, and curves, then mask them based on that using one of the image segmentation approaches (Semantic Segmentation and Instance Segmentation). As shown in Fig. 12, mask R-CNN has three branches, one for class label object, one for bounding offset, and one for masking these outputs into different colors.
- Mask R-CNN has a lot of pros and cons; most pros like it consider simple, fast, efficient, and flexible with good performance compared with other ANN techniques like R-CNN and Fast R-CNN [14].

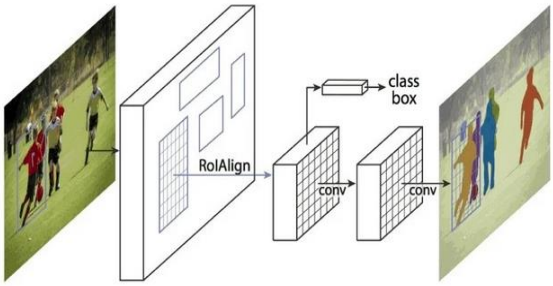

Fig 12. Mask R-CNN Architecture

II.VI DIMENSIONALITY REDUCTION BY FEATURE SELECTION AND FEATURE EXTRACTION

It is the process of reducing the number of required features to analyze and identify the objects on images by using multiple parameters and dividing them into two ways [15].
First is feature selection [16], which uses the most related and proper variables from the given dataset using different approaches such as correlation, forward selection, and selecting the best key K value.
The second is Feature extraction, which uses a small set of variables, including input variables; there are a lot of techniques for this stage, and we will use most of them in this paper and make some comparisons between them and choose the techniques that provide the best results, below some of these techniques:

I. PCA (Principal Component Analysis) [17]
II. LDA (Linear Discriminant Analysis) [18]
III. HoG (Histogram of oriented gradients)
IV. CTD (discrete cosine transform domain)
V. Scale-Invariant Feature Transform (SIFT)
VI. Speeded-Up Robust Feature (SURF) [19]
VII. Convolutional neural network (CNN)

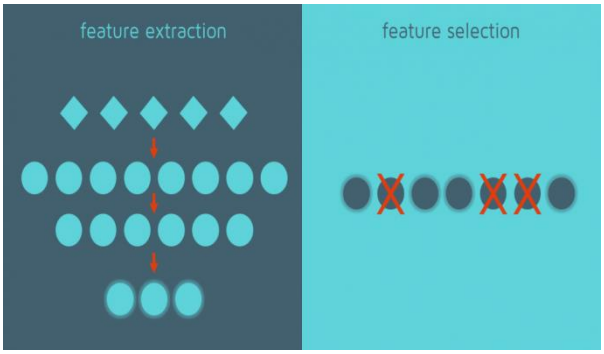

Fig 13. Feature Extraction vs. Feature Selection

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

II.VII SIMILARITY RETRIEVAL (SR)

Based on how colors are distributed in the image, the pixels will be divided into groups based on the continuous Region of the color. Some pixels will be coherent, and some are not [20]. This is how retrieving the image will be classified based on this and do clustering based structural index matrix SSIM methodology to do image similarity retrieval process**.**

II.VIII AUTOMATIC TARGET RECOGNITION

It is a system that immediately recognizes the objects based on input sensors used in industries like manufacturing, military "Radar," and private sectors such as CCTV and traffic monitoring systems. The ultimate goal is to enhance the capabilities and increase the speed of object recognition to provide the system's input with very known objects based on pre-trained datasets to reach the system's goal.

- Possible military applications include a simple identification system used in other applications such as uncrewed aerial vehicles, cruise missiles, CCTV for border security, and safety systems to identify objects or people.

III. METHODOLOGY

A description of how the different parts/stages of the Proposed system are implemented. As shown in the flowchart of how the other parts of the system interrelate Fig. 14, First Stage is feature extraction, divided into two parts one is the training part and the second is the evaluation part.

Both parts will divide the images into two groups, one for good & bad quality images and one for good & bad content, then send the suitable images to the second classification stage, which is the evaluation part.

The first stage is the feature selection stage, which will select and pickup required features of existing objects in images using different algorithms:

I. PCA (Principal Component Analysis)
II. LDA (Linear Discriminant Analysis)
III. HoG (Histogram of oriented gradients)
IV. DCT (discrete cosine transform domain) [21]
V. Scale-Invariant Feature Transform (SIFT)
VI. Speeded Up Robust Feature (SURF)
VII. Convolutional neural network (CNN)

- The best algorithm that gives the best results will be chosen for the next stage.
- The second stage, the classification Stage, will use the output of the first stage using good images to do more classification processes using multiple image classification algorithms and techniques such as support vector machine SVM, random forest RF, and ensemble. The next stage will choose the algorithm that gives the best results using Accuracy, Precision and Recall evaluation.
- The Third Stage is the cluster stage, which uses different techniques to divide images into groups using K-means and LBP or any unsupervised machine learning algorithms. The best output results will move to the next stage.
- The fourth stage is the detection stage using image segmentation techniques that use different techniques such as KNN, K-Means, and ANN [22] algorithms using Mask R-CNN, and the best output results will move to the next stage.
- The final stage is where one image is selected from a cluster of similar images and then passed to the ATR system for suitable and quick object recognition and tracking [23].

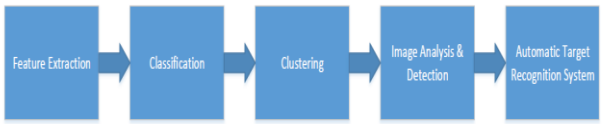

Fig 14. Image Processing Proposed System Flow

III.I FEATURE EXTRACTION

Many methods exist for feature extraction, and this paper will focus on three methods: HOG, DCT, and CNN. The best results of these methods will be chosen to move forward to the next stage.

HOG mainly uses gradients to provide information about the image's content, especially for edges and corners, which are more fit for object detection.

DCT is mainly used to convert spatial information to frequency information that provides more information about images' quality and divides images based on their frequency parameters.

CNN is mainly used for image classification. It is suitable for feature image extraction since it reduces the required number of parameters without affecting the image quality with high accuracy. Network layers are trained for a massive set of images suitable for other tasks, especially object recognition.

III.II IMAGE CLASSIFICATION

There are many algorithms & techniques used for image classification, and this paper will focus on two algorithms: SVM, RF, and ensemble learning model that combine different algorithms (SVM, RF, KNN, and DT).

SVM, mainly used for the classification and regression of images, can reduce segmentation errors on moving objects and solve the issue of overfitting in a vast image dataset with large features.

RF is mainly used to classify extensive data, with two main parameters: the number of trees and the number of features when talking about image classification.

The ensemble learning model [24] can be used with different techniques like bagging, stacking, and boosting, and use a mix of algorithms such as SVM, RF, and KNN and provide one layer with a meta classifier.

This stage will use the algorithm or model that provides the best results and pass it to the next stage.

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

III.III Image Clustering

K-Means [25] is the process of grouping similar images and each other's to guarantee non-overlapping between groups. It can be used in image clustering segmentation and classification using the MNIST dataset. K-means can also be used in image segmentation to cluster images into related groups correctly. The K-means process begins to extract the required features and cluster them within groups or partitions with similar features for k-clusters based on the centroid for each k value.

LBP [26] (Local Binary Patterns) is mainly used in the feature extraction process "texture" to do labeling for input image pixels by using threshold techniques, and the result will be a binary number. This means we can then do a clustering approach using k-means to provide the suitable and required feature in the image. Moreover, LBP uses local-based feature representation, especially in facial image analysis, and includes processes in object tracking methodologies such as face detection, face representation, facial analysis, classification, and face recognition.

III.IV Image Analysis & detection Using segmentation techniques

As described in section # 2, many image segmentation techniques can be used for images. This stage will use two techniques: clustering-based image segmentation using K-Means and ANN-based image segmentation using the Mask R-CNN method. The first process will use mask R-CNN to detect the object in the image and then use K-Means to do image segmentation to get the ultimate results that can easily distinguish the object to provide it to the next stage, which is an automatic recognition system to minimize the time and computation when looking for a specific object into massive datasets and especially on the non-static environment such as for CCTV system [28].

III.V Object recognition & Automatic target recognition system ATR

The main idea behind this approach is to feed the ATR system with the ability to detect the object in real-time and with minimum time and less computational resources, Using CNN techniques based on VGG[29], Rest Net[30], and ImageNet[31] methods to achieve the goal behind this proposed methodology. The COCO dataset [32] will be used as a test benchmark of the final model results.

IV. Evaluation of the Proposed System

The proposed system will be evaluated for each stage, choosing the best results of different use techniques and then putting all evaluation stages together to assess the whole system once all evaluation values are combined.
The evaluation criteria will be built on open measures values of accuracy, precision, and recall for retrieval images.

Precision: Describes how many of the retrieved images should be retrieved. *(4.1)*

Recall: Describes how many of the images that should be retrieved are retrieved. *(4.2)*

Accuracy: Describes how many classifications are out of all classifications made. *(4.3)*

Values consist of the following values:
True Positive TP; False Positive FP; True Negative TN; False Negative FN.

$$Precision = \frac{true\ positives}{true\ positives + false\ positives}$$

$$Recall = \frac{true\ positives}{true\ positives + false\ negatives}$$

$$Accuracy = \frac{true\ positives + true\ negatives}{all\ samples}$$

- This paper will use a COCO data set containing 328,000 images with almost 2.5 million instances and 80 classes. We use structural similarity index matrix SSIM techniques to filter authentic images from degenerate images to achieve the best quality image compared with the original image by measuring the similarity between images.
So, depending on this methodology, we below show a sample as shown in Fig. 15. Our techniques can differentiate between authentic and original images and degenerated images for a flower with a percentage of 75 %.- SSIM has a range value between [0-1], with best values that are near one and wrong values near zero, and we choose 0.75 consider from better quality images and put them in one category that will be used for both training and test or evaluation process that will be used in different stages. SSIM can also be used for image compression, image restoration, and pattern recognition to simulate human perception.

Fig. 15: An example of a degraded image, the authentic one is seen on the left, and the degraded image is blurred on the right

This task has been processed into two-level to measure the model's accuracy, and one includes original and non-original images, as shown below in fig. 16—second, use

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

only original and authentic images. After that, the final images in the specified category will be split into a training set and testing set of 80% and 20%, respectively, including both original and degenerated images to do the evaluation.

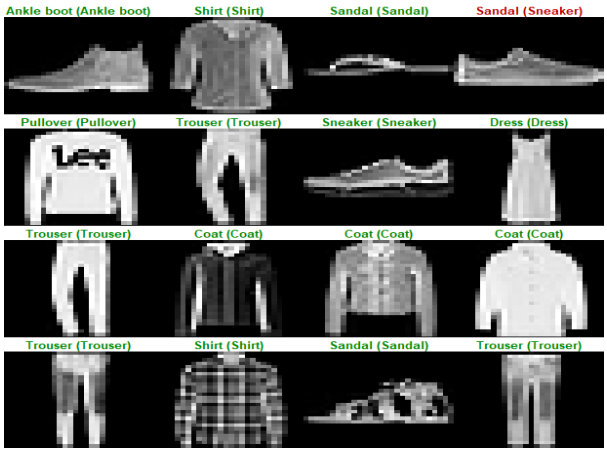

Fig. 16: Examples of similar images belonging to the same cluster.

V. EXPERIMENT SETUP & RESULTS

The simulation of the proposed system was built using different tools: MATLAB, RStudio, and Golab, including computer vision and neural networks and machine learning packages like TensorFlow and Kiras. The primary goal of this paper is to create a model that provides the best results for detecting and recognizing an object in images with minimum time and less computational resources than used in an automatic recognition system which in this paper will be a CCTV system.

- Experiments were performed on a single node laptop, four cores, Microsoft Windows 10, Core i7 processor, 1.8GHz speed, 16 GB memory, and 500 GB HDD and google lab "Golab." Different data sets and algorithms are used to build the model. The data sets used are COCO and VCC16, which contains many images; we chose a sample of "3000 images," including a specific object such as flowers and animal "Cat & Dogs," to experiment. The algorithms are HOG, DCT, R-CNN, SVM, RF, LBP, Learners, and Mask R-CNN. As shown in Table 1-4, the results using separated algorithms compared to our proposed model provide the best results.

Table 1: Features Extraction Stage

Type	Total Images	HOG	DCT	RCNN
True Positive	3000	2245	2345	2323
False Positive	3000	304	203	225
False Negative	3000	710	620	462
True Negative	3000	259	168	100
Accuracy	100	83	84	93
Precision	100	88	90	91
Recall	100	76	79	83

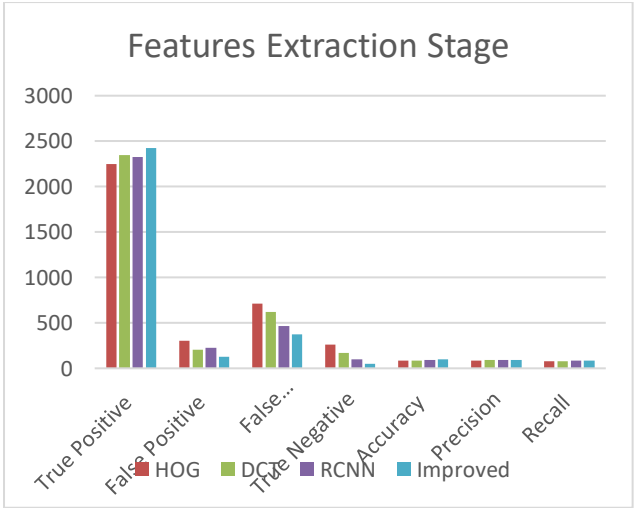

Fig. 17: Feature Extraction Stage Results

As shown in Table 1, describe the results of stage one in the proposed model. The R-CNN model provides the best accuracy for identifying the required features at 93 % compared to DCT with 84% and HOG at 83%, respectively. So, based on our adopted methodology will choose the output of R-CNN images that will move forward to the next stage.

Table 2: Classification Stage

Type	Accuracy	Precision	Recall
SVM	97.7	88.9	88.9
RF	96.9	84.5	84.8
KNN	97.2	86	88.2
Learners	98.8	89.6	89.5

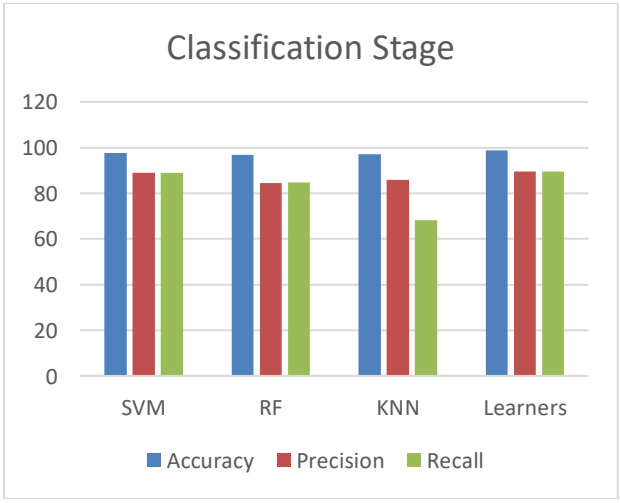

Fig. 18: Classification Stage Results

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

As shown in Table 2 describe the results of stage one in the proposed model.
The learner model mixes more than one algorithm and provides the best accuracy for image classification with 98.8 % compared to SVM at 97.7%, KNN at 97.2%, and RF at 96.9 %, respectively. So, based on our adopted methodology will choose the output of Learner model images that will move forward to the next stage.

Table 3: Clustering Stage using pretrained images data set VGG16

Type	Accuracy	Precision	Recall
K-means	96	98	92
KNN	96.5	98.4	93
DT	97.3	99.4	94
Logistic Regression	98	99.4	96

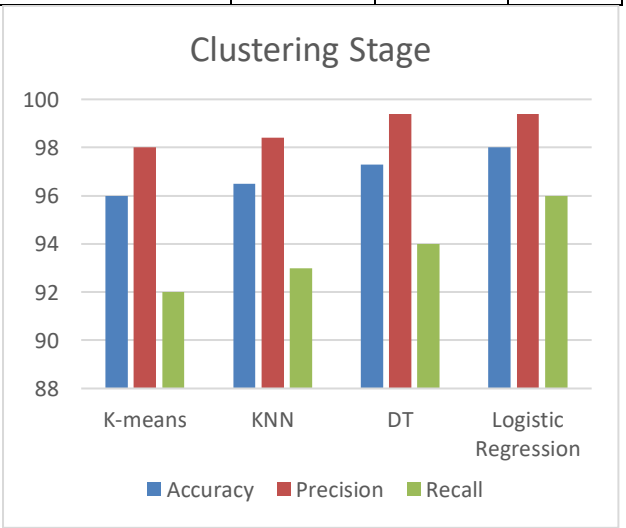

Fig. 19: Clustering Stage Results

As shown in Table 3 describe the results of stage one in the proposed model.
Logistic regression provides the best accuracy results for image clustering with 98 % compared to DT 97.3%, KNN 96.5%, and K-means 96 %, respectively. So, based on our adopted methodology will choose the output of DT images that will move forward to the next stage.

Table 4: Image Segmentation for Object detection & recognition:

Object	Accuracy	Precision	Recall
Cat	78	77.8	93
Dog	74	77	88
Person	82	82	95
Flower	79	80	90
Airplane	76	80	92

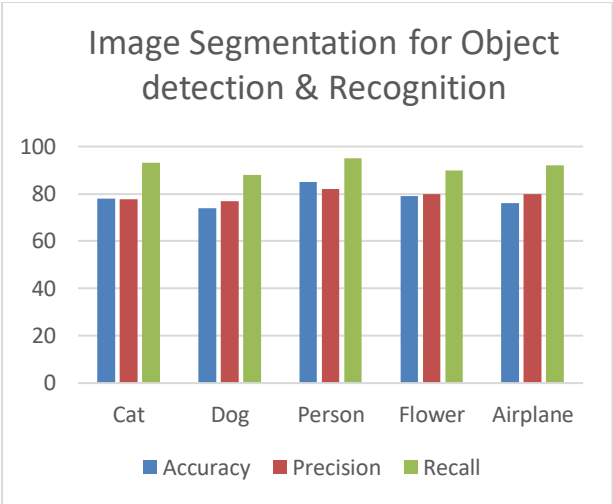

Fig. 20: Image Segmentation Stage Results

As shown in Table 4 describe the results of stage one in the proposed model.
Image segmentation-based artificial intelligence techniques using Mask R-CNN provide the best accuracy results for most objects from sample data set images ranging between 74 – 82 %. So, based on our adopted methodology will choose the output of the model images that will move forward to the next stage.

Table 5: Final Output to be provided to ART – CCTV.

Object	Accuracy	Precision	Recall
Cat	83	84	98
Dog	79	83	93
Person	90	87	99
Flower	84	85	95
Airplane	83	85	97

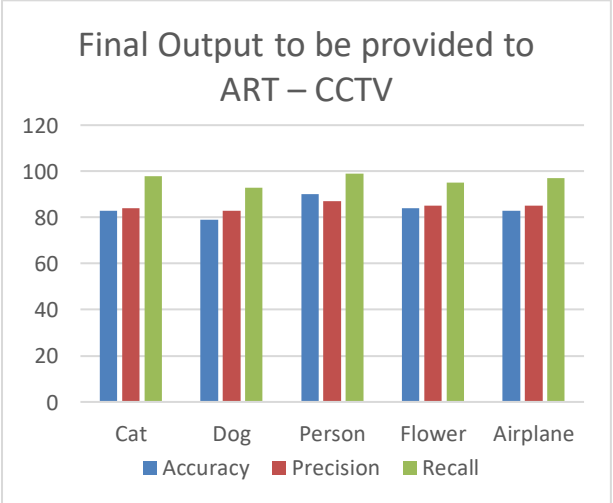

Fig. 20: ATR – CCTV Object Detection Results

As shown in Table 5 describe the results of stage one in the proposed model.
Object Detection & recognition is the last stage in the prosed model that provides better accuracy after passing

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

images to multiple filtrations. Detection accuracy varies between 79 – 90 % and is considered more accurate and efficient than filtration only using one technique, as shown in Table 4. Moreover, as shown in Fig 22 & Fig. 23, the ATR system could easily differentiate between objects compared with Fig. 21 using a segmentation approach based on the proposed model.

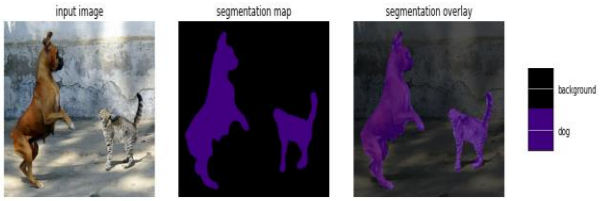

Fig. 21: Image Segmentation Result using one technique

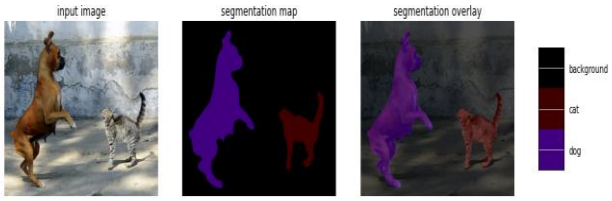

Fig. 22: Proposed System Results

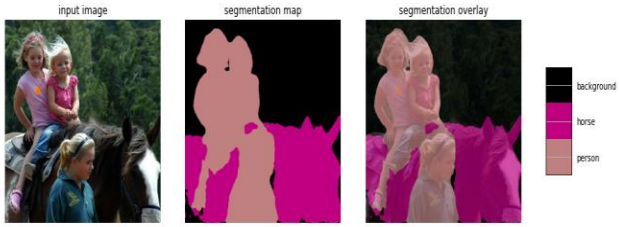

Fig. 23: Proposed System Results

VI. Conclusion & Future Work

Using a digital image processing approach with different techniques to build another method of object detection and recognition used in the computer vision approach to be part of an automatic recognition system such as a surveillance system CCTV is the aim of this paper.

The proposed model uses more than one stage to filter the images before entering the CCTV system to provide the best results with high accuracy and fewer computational resources.

Stage one of the proposed model's built-based feature extraction techniques using HOG, DCT, and R-CNN for a sample of COCO data set and providing the highest accuracy using R-CNN techniques, then passing these images to the second stage.

Stage two uses classification techniques like random forest RF, KNN, SVM, and ensemble learner to provide the highest accuracy using the ensemble model to pass for the third stage.

Stage three uses image clustering techniques like LR, DT, KNN, and K-means, which provide the highest accuracy using LR that pass to the final stage. Stage four and the final one used image segmentation techniques like K-means and Mask R-CNN; object detection and recognition accuracy in this stage using artificial neural network Mask R-CNN reached 82 %, and when the final image set passed to the ATC-CCTV system, the accuracy increased and reach up to 90 %.

In future works, we could improve our model and build a prototype for the proposed model to demonstrate the model's feasibility in analyzing CCTV images.

Acknowledgments

An acknowledgment section may be presented after the conclusion if desired.

References

[1] Aparna Akula, Arshdeep Singh, Ripul Ghosh, Satish Kumar, and Hk Sardana, Target Recognition in Infrared Imagery Using Convolutional Neural Network, 2017.

[2] Eren Golge. How does feature extraction work on images? URLhttps://www.quora.com/profile /Eren-Golge/Machine-Learning/How-does-feature-extraction-work-on-images.

[3] F. Suard, A. Rakotomamonjy, and A. Bensrhair. Pedestrian detection using infrared images and histograms of oriented gradients. In IEEE Conference on Intelligent Vehicles, 2006.

[4] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman. Return of the devil in the details: Delving deep into convolutional nets. At British Machine Vision Conference, 2014.

[5] MathWorks. Support vector machines for binary classification.URL https://se.mathworks.com/help/stats/ support-vector-machines-for-binary-classification.html.

[6] Xia, J.; Ghamisi, P.; Yokoya, N.; Iwasaki, A. Random forest ensembles and extended multiextinction profiles for hyperspectral image classification, 2017.

[7] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. Trans, 2004.

[8] David G. Lowe, Distinctive image features from scale-invariant key points," Int. J. Computer Vision, 2004.

[9] Burhan Duman, Ahmet Ali Süzen, A Study on Deep Learning Based Classification of Flower Images, International Journal of Advanced Networking and Applications (IJANA), 2022.

[10] He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017.MaskR-CNN,12pp. http://arxiv.org/pdf/1703.06870v3.

[11] M Sharif, M Raza, S Mohsin, JH Shah, Microscopic Feature Extraction Method, International Journal of Advanced Networking and Application (IJANA), 2013.

Volume: 14 Issue: 04 Pages: 5499-5508(2023) ISSN: 0975-0290

[12] Convolutional neural networks (lenet). URL http://deeplearning.net/ tutorial/lenet.html, 2020.
[13] Ren, S., K. He, R. Girshick, and J. Sun, Faster r-CNN: Towards real-time object detection with region proposal networks. In C. Cortes, N. D. Lawrence, D. D. Lee, and R. Garnett (Eds.), Advances in Neural Information Processing Systems , 2015.
[14] Girshick R. Fast r-CNN. In: Proceedings of the IEEE international conference on computer vision; 2015.
[15] Merentitis, A.; Debes, C.; Heremans, R. Ensemble learning in hyperspectral image classification: Toward selecting a favorable bias-variance tradeoff. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 2014.
[16] Ling C., Bolun C., and Yixin C., Image Feature Selection Based on Ant Colony Optimization, 2011.
[17] Uddin, M.P.; Mamun, MA; Hossain, M.A. PCA-based feature reduction for hyperspectral remote sensing image classification. IETE Tech. Rev. 2021
[18] A. Riddhi, Vyas D., and Shah S. Comparison of PCA and LDA techniques for face recognition feature-based extraction with accuracy enhancement. IRJET, 2017.
[19] R. Hendaoui, M. Abdellaoui, and A. Douik, "Synthesis of Spatio-temporal interest point detectors: Harris 3D, MoSIFT and SURF-MHI," in Proc. 1st Int. Conf. Adv. Technol. Signal Image Process, 2014.
[20] Greg Pass, Ramin Zabih, and Justin Miller. Comparing images using color coherence vectors. In Proceedings of the Fourth ACM International Conference on Multimedia, 1996.
[21] MathWorks. Discrete cosine transform, URL https://se.mathworks.com/help/images/discrete-cosine-transform.html, 2020.
[22] Michael A. Nielsen. Neural Networks and Deep Learning. Determination Press, 2015.
[23] Antoine d' Acremont, Ronan Fablet, Alexandre Baussard, and Guillaume Quin, Cnn-based target recognition and identification for infrared imaging in defense systems," Sensors, vol. 19, 2019.
[24] Dominik Müller, Iñaki Soto-Rey and Frank Kramer, An Analysis on Ensemble Learning optimized Medical Image Classification with Deep Convolutional Neural Networks, IT-Infrastructure for Translational Medical Research, 2022.
[25] Lloyd, Stuart, least squares quantization in PCM, In IEEE transactions on information theory, 1982.
[26] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns IEEE Trans. Pattern Anal. Mach. Intell., 24 (7), 2002.
[27] Huang, d., Shan, C., Ardabilian, M., Chen, L.: Local binary patterns and its application to facial image analysis: A survey. IEEE Transactions on Systems, Man, and Cybernetics, 2011.
[28] Hanae Moussaoui, Mohamed Benslimane and Nabil El Akkad, Image Segmentation Approach Based on Hybridization Between K-Means and Mask R-CNN, Springer Singapore, 2020.
[29] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2016.
[30] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016.
[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in Computer Vision and Pattern Recognition, IEEE Conference on, 2009.
[32] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll'ar, C. L. Zitnick, Microsoft coco: Common objects in context, in European conference on computer vision, Springer, 2014.

Author Biography

Mr. Adel Hassan received his undergraduate Degree in Electrical Engineering from Bierzeit University, Palestine, in 2000. He received his Post- Graduate Degree in Computer Science, M.Sc.(CS), from Arab American University, Palestine, in 2019 and a candidate for Ph.D. in Information Technology Engineering in 2023. Presently he is working at the National bank of Palestine as chief technology officer (CTO).

Muath Sabha is Assistant Professor in Computer Engineering who works at Arab American University as a Professor and Researcher. He has developed and led development teams for several curricula, including Computer Graphics, Educational Technology, Computer Engineering, and Multimedia Technology programs. He was a consultant for many universities and the Palestinian Ministry of Higher Education for Program Development and Evaluation. He trained in many courses in interactive multimedia production, especially in education. Muath has research interests in multimedia, multimedia in education, multimedia in cultural heritage, computer vision, augmented reality, and computer graphics. He is a reviewer in many conferences and journals. Muath has been awarded many scholarships, including his Ph.D., Master's Degree, and Partially his BSc. Muath holds a Ph.D. in Computer Graphics from the University of Leuven (former name KULeuven), Belgium.