# Motion Detection and Tracking of Classified Objects with Intelligent Systems

**Aydın Eresen*. Nevrez İmamoğlu.***
**Mehmet Önder Efe***

**Dept. of Electrical and Electronics Engineering, TOBB Economics and Technology University, Söğütözü, Ankara Turkey (Tel: +90-312-292 4064; e-mail: {aeresen,nimamoglu,onderefe}@etu.edu.tr).*

**Abstract:** In this paper, detecting and tracking of a car is aimed using a stationary camera system. Background subtraction is used to detect motion and Kalman filter is used for tracking of a moving car. Classification of the car is accomplished utilizing Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs). In using SVMs and ANNs, some features should be extracted from Region of Interest (RoI). Before the extraction, image enhancement methods are used and then by using Discrete Wavelet Transform (DWT), the features are represented in different frequency scales. The current work compares the performances of ANNs trained via Levenberg-Marquardt optimization technique, Least Squares Support Vector Classification (LS-SVC) and υ Support Vector Classification (υ - SVC).

*Keywords*: Neural Networks, Support Vector Machines, Image Segmentation, Image Recognition

## 1. INTRODUCTION

Detection and recognition of an object is of crucial importance especially in unmanned surveillance systems. These extend from fixed cameras as well as platforms that move in the 3D space. An unmanned aerial vehicle (UAV) is a good example, where the camera is moving, the scene is changing and object is moving within the scene. Detecting an object and tracking of it is a state-of-the-art research problem addressed by several researchers so far. Clearly the video based systems enjoy the techniques of motion detection algorithms heavily. In the literature many algorithms are mentioned, background subtraction methods (Piccardi, 2004), using peripheral and foveal vision (Gould, 2007), hybrid algorithm and wavelet based detection (Töreyin, 2003), an improved algorithm on adaptive gaussian mixture model (Zivkovic, 2004), nonparametric kernel density estimation (Elgammal, 2002) are just no name a few. Among these algorithms, simple background subtraction method has less memory requirement and computational complexity compared to its alternatives.

Almost all video based applications require an image enhancement stage. There are several works about image enhancement such as image smoothing (Andrade, 2004), denoising and soft-thresholding (Donoho, 2004; Gonzalez, 2008) and fusing (Irani, 1993). Image enhancement is a critically important stage influencing the performance of recognition and classification mechanisms.

Classification of the RoI is another problem in the image processing applications. The main part of the classification problem is to choose features from RoI. Feature selection varies according to the environment and characteristics of the problem in hand (Gonzalez, 2008; Nixon, 2008; Borchani, 1997). Instead of extracting features of only RoI, the features spread over all frequencies and scales obtained from DWT are extracted to attain more accurate classification (Walker, 1999; Sarlashkar, 1998; Jamarani, 2005). There are many classification methods in literature, e.g. back-propagation (Jamarani, 2005), bayesian classifiers and neural networks (Zhang, 2000; Russel, 2003), artificial neural networks (Haykin, 1999), υ -support vector classifiers (Chen, 1992) and least squares support vector classifiers (Suykens, 2002). Among these approaches, ANNs and SVMs have become frequently used methods as the former has very powerful learning algorithms to learn input-output data and the latter converts the learning problem into an optimization problem and minimized the upper bound of the empirical risk. This paper considers classification with Levenberg-Marquardt trained ANNs, υ-SVCs and LS-SVCs (Lourakis, 2005; Chen, 1992; Suykens, 2002). Tracking will be performed if the classification of the moving car succeeds. In the literature, tracking is generally achieved by using a Kalman filter (Russel, 2003; Niu, 2003) which will be the method adopted here.

The paper is organized as follows: Section 2 explains the segmentation of RoIs. Section 3 expresses the features selected from RoI. Section 4 demonstrates the implementation of intelligent systems stated in this paper. Section 5 demonstrates the results of the classified system. Concluding remarks are given in Section 6.

## 2. SEGMENTATION OF RoIs

### *2.1 Detection of RoIs with Motion*

Motion detection in video sequences is a prime problem in tracking applications. There are many different methods for motion detection as stated above. All of them have advantages and disadvantages. Simple Background Subtraction (SBS) is chosen for detecting motion to segment objects which has possibility of being a car. SBS method can be explained as follows:

$$F_{result} = \begin{cases} 1, & I_t(x,y) - I_{t-1}(x,y) > Threshold \\ 0, & I_t(x,y) - I_{t-1}(x,y) \leq Threshold \end{cases} \tag{1}$$

Application of the thresholding in (1) results in a binary image in which moving areas are white, others are black. In our application, instead of using binary image, the result is converted to grayscale image. The area of the motion represents the RoI that will be used for classification whether the object within the region is a car or not.

### *2.2 Detection of RoIs without Motion*

If there is no motion in video sequences, to determine the RoI, object detection algorithms are used. These objects will be the RoIs which have the possibility of being a car, all objects are classified whether they are cars or not. A flow chart of the proposed system is as follows,

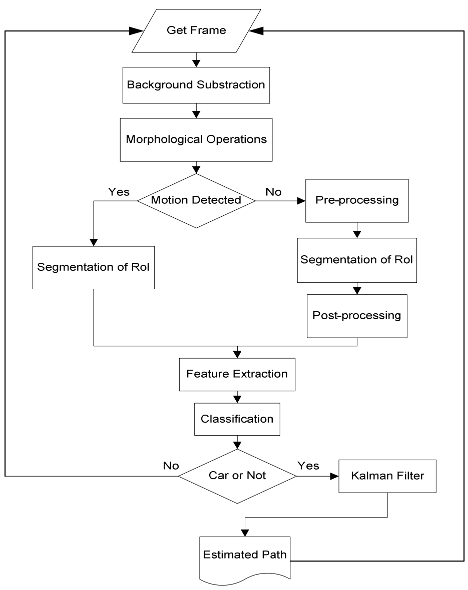


Fig. 1: Flow Chart of the Proposed System

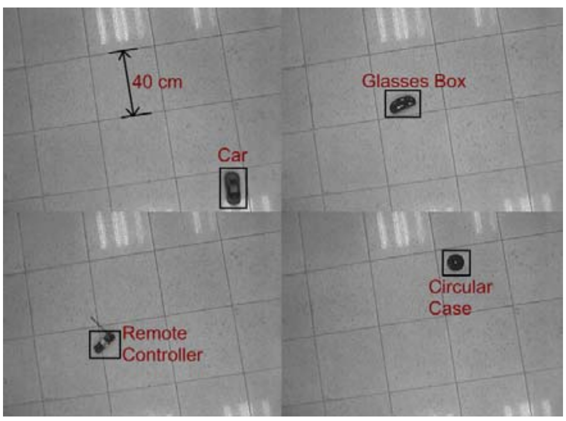


Fig. 2: Obtained frames from the system

Fig. 2 shows the various frames with different objects chosen for classification. Segmentation results of these images shown Fig. 3 are used to extract features for classification.

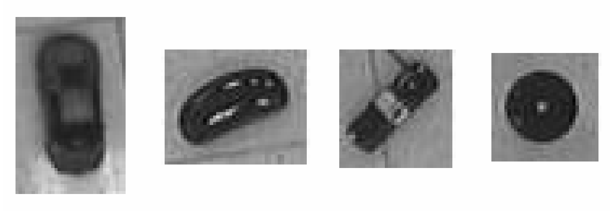

Fig. 3: Segmented objects from different video frames

## 3. FEATURE SELECTION FOR CLASSIFICATION

All classification methods need features which are extracted from the data. In this paper, the features are extracted from the RoIs segmented from whole image.

### *3.1 Selected Features*

In this paper, first and second order statistics of the image are used. The expressions used to obtain features proposed by Borchani and Stamon (1997) are given below. For the first order statistics from image histogram we have (2)-(5).

$$M = \mu = \sum_{i=1}^{n} i h(i) \tag{2}$$

$$SD = \sum_{i=1}^{n} (i-\mu)^2 h(i) \tag{3}$$

$$TM = \sum_{i=1}^{n} (i-\mu)^3 h(i) \tag{4}$$

$$E_1 = -\sum_{i=1}^{n} h(i) \log(h(i)) \tag{5}$$

where *h(i)* is the image histogram, *M* is the mean of the histogram, *SD* is the standard deviation of histogram, *TM* is third moment of the histogram, $E_1$ is the first order entropy.

Similarly the second order statistics from an image histogram are obtained as given by (6)-(16).

$$H = \frac{1}{N_c^2} \sum_i \sum_j (M_c(i,j))^2 \tag{6}$$

$$Cont = \frac{1}{N_c (L-1)^2} \sum_{k=0}^{L-1} k^2 \sum_{|i-j|=k} M_c(i,j) \tag{7}$$

$$E_2 = 1 - \frac{1}{N_c (L-1)^2} \sum_i \sum_j M_c(i,j) \log(M_c(i,j)) \tag{8}$$

$$Corr = \frac{1}{N_c \sigma_x \sigma_y} \left| \sum_i \sum_j (i-m_x)(j-m_y) M_c(i,j) \right| \tag{9}$$

$$m_x = \frac{1}{N_c} \sum_i \sum_j i M_c(i,j) \tag{10}$$

$$m_y = \frac{1}{N_c} \sum_i \sum_j j M_c(i,j) \tag{11}$$

$$\sigma_x^2 = \frac{1}{N_c} \sum_i \sum_j (i-m_x)^2 M_c(i,j) \tag{12}$$

$$\sigma_y^2 = \frac{1}{N_c} \sum_i \sum_j (i-m_y)^2 M_c(i,j) \tag{13}$$

$$LH = \frac{1}{N_c} \sum_i \sum_j \frac{1}{1+(i-j)^2} M_c(i,j) \tag{14}$$

$$D = \frac{1}{N_c} \sum_i M_c(i,i) \tag{15}$$

$$U = \frac{1}{N_c^2} \sum_i M_c(i,i)^2 \tag{16}$$

where $M_c$ is gray level co-occurrence matrices and $N_c$ is the sum of the elements of $M_c$, $H$ is homogeneity, *Cont* is contrast, $E_2$ second order entropy, *Corr* is the correlation, *LH* is the local homogeneity, $D$ is the directionality, and $U$ is the uniformity.

### *3.2 Wavelet Transform*

Wavelet Transform (WT) is a useful tool to represent signal in different frequency scales. WT decomposes a signal into two sub-signals which can be called the approximate (low frequency component) and the detail (high frequency component) sub-signals (Walker, 1999). These sub-signals have the half size of the original signal. The proposed algorithm uses WT in 2-D to represent RoI with different frequency scales. The following expressions for the WT equations are stated by Walker (1999). The most simple 1-D WT method is Haar Transform. Approximate signal is $a^1 = a_1, a_2, \ldots, a_{N/2}$, and detail signal is $d^1 = d_1, d_2, \ldots, d_{N/2}$. First level Haar transform can be described as follows (Walker, 1999)

$$Haar(g) = (a^1 \mid d^1) \rightarrow a_n = \frac{f_{2n-1} + f_{2n}}{\sqrt{2}}, d_n = \frac{f_{2n-1} - f_{2n}}{\sqrt{2}} \quad (17)$$

Although Haar wavelet is the simplest one, alternative forms may yield better energy preservation results such as Daub4, which is applied as Haar Transform but the number of samples is four instead of two. The coefficients used to calculate $a$ and $d$ are as follows,

$$\alpha_1 = \frac{1+\sqrt{3}}{4\sqrt{2}}, \alpha_2 = \frac{3+\sqrt{3}}{4\sqrt{2}}, \alpha_3 = \frac{3-\sqrt{3}}{4\sqrt{2}}, \alpha_4 = \frac{1-\sqrt{3}}{4\sqrt{2}} \quad (18)$$

$$\beta_1 = \frac{1-\sqrt{3}}{4\sqrt{2}}, \beta_2 = \frac{\sqrt{3}-3}{4\sqrt{2}}, \beta_3 = \frac{3+\sqrt{3}}{4\sqrt{2}}, \beta_4 = \frac{-1-\sqrt{3}}{4\sqrt{2}} \quad (19)$$

where $\alpha$ is used to calculate the approximate signal, $\beta$ and is used to calculate the detail signal.

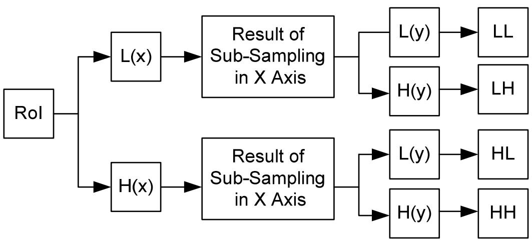


Fig. 4: 2-D Wavelet Transform

2-D WT can be applied by using two 1-D WT in both axes. Using 2-D WT on RoI, we obtain 4 sub-images which represent different frequency components. Fig. 4 summarizes how 2-D WT works.

## 4. INTELLIGENT CLASSIFIERS

Intelligent systems are used in various applications such as automatic control, prediction, speech processing, object recognition, image enhancement, and classification. In this study, ANNs and SVMs are used for classification and recognition as mentioned before. The system receives 55 input vector, features extracted from the object, and outputs a scalar value determining whether the detected object is a car or not.

### *4.1 Artificial Neural Networks*

ANNs are nonlinear models composed of weights in between interconnected components called neurons. The proposed ANN model is trained by using Levenberg-Marquardt optimization method which is a second order optimization algorithm. Using Newton and Gradient Descent methods, Levenberg-Marquardt technique minimizes the cost function $J$ in (20) while avoiding getting stuck to local minima.

$$J = \frac{\partial f(p)}{\partial p} \quad (20)$$

where $p$ is the parameter vector and $f$ is the vector of sample errors. For a thorough treatment one should refer to Lourakis and Argyros (2005) and Hagan and Menhaj (1994).

### *4.2 Support Vector Machines*

A SVM is an effective pattern classification method proposed by Boser et al (1992). Traditional classification methods try to minimize empirical training error but the purpose of the SVM is to minimize upper bound of the generalization error between nearest data and separating line, and generalization performance of the SVMs is better than ANNs. During the process of the minimization to reach the global minimum, SVM prevents the system to stop at a local minimum.

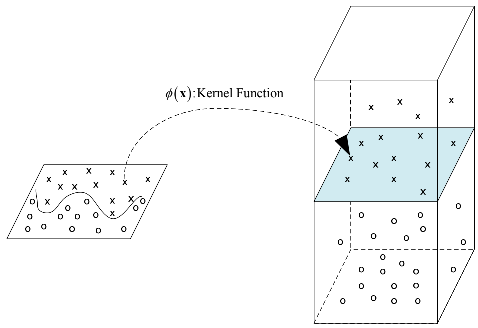


Fig. 5: Kernel function maps input space to high dimensional feature space

SVM is originally a linear classification method admitting linearly separable training data. Kernel trick is introduced to tackle the classification problems requiring nonlinear decision boundaries. Kernel trick can be considered as a transformation. Data set is mapped to a high dimensional feature space (See Fig. 5), in which the transformed data becomes linearly separable. The performance of the SVM classifier is dependent upon the selected kernel function and unfortunately a guide is unavailable in choosing a particular kernel. Due to its popularity, in this paper, Radial Basis

Function (RBF) kernel is used for nonlinear case. RBF kernel function is given at (21) (Burges, 1998).

$$K\left(x_i,x_j\right)=\exp\left(-\frac{\left\|x_i-x_j\right\|^2}{2\sigma^2}\right) \tag{21}$$

where $x_i$ and $x_j$ vectors are input data and explained in nonlinear classification with SVM.

### *4.2.1 Linear Classification with SVM*

Linear SVC is known also as Maximal Margin Classifier (MMC). MMC is the most practical SVM to implement, yet the classification errors with MMC are typically high as it requires the data to be separable somehow. This is indeed a disadvantage as many real time measurements in the neighbourhood of the decision region may lie on incorrect side causing some error in the classification result. Soft Margin Classifiers (SMCs) is proposed to alleviate this disadvantage though penalizing the misclassifications appropriately (Chen et al, 2005). In SMC theory, smaller and acceptable levels of errors are achievable while creating a suitable separating plane.

In SVM theory, firstly primal form of the optimization problem is constructed and then dual form of the minimization problem is obtained using the method of Lagrange multipliers. Once the dual representation is obtained, the problem becomes a quadratic optimization problem, whose solution is straightforward. As discussed in Chen et al (2005), linear SVM can be described as below. Let $S$ be the set of input-output pairs.

$$S=\left\{\left(x_1,y_1\right),\ldots,\left(x_k,y_k\right)\right\}, x\in\mathbb{R}^n, y\in\left(-1,+1\right) \tag{22}$$

where $\boldsymbol{x}$ denotes the inputs, $\boldsymbol{y}$ stands for the target output (class). The separating hyperplane is described by (23),

$$\left\langle w,x\right\rangle+b=0 \tag{23}$$

where $\boldsymbol{w}$ is weight vector and $b$ is the bias value. The primal form of the classifier is given in (24).

$$f(x)=\operatorname{sgn}\left(\left\langle w,x\right\rangle+b\right) \tag{24}$$

As shown in Fig. 6, the distance between $<\boldsymbol{w},x^+>+b=+1$ and $<\boldsymbol{w},x^->+b=-1$ is called the margin denoted by $m$, where $x^+$ is positive class data and $x^-$ is negative class data. The maximum value of $m$ is sought for the optimal hyperplane, and we have

$$m=\frac{2}{\|w\|} \tag{25}$$

so that the norm of $\boldsymbol{w}$ should be minimized. The primal form of the optimization problem is given in (26).

$$\min_{w,b}\frac{1}{2}\|w\|^2 \quad st \quad y_i\left(\left\langle w,x_i\right\rangle+b\right)\geq 1,\ i=1,\ldots,m \tag{26}$$

and by introducing the Lagrange multipliers, the dual form of the optimization problem is obtained as in (27).

$$\max_{w,b}\sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i,j}^{m}\alpha_i\alpha_j y_i y_j\left\langle x_i,x_j\right\rangle \quad st \quad \alpha_i\geq 0, i=1,\ldots,m, and \sum_{i=1}^{m}\alpha_i y_i=0 \tag{27}$$

where $\alpha_i$ denotes $i$-th Lagrange multiplier. $\alpha_i$, ($i$=1...$m$) are calculated using a quadratic problem solver, and then hyperplane decision function is obtained as

$$f(x,\alpha^*,b^*)=\sum_{i=1}^{N}y_i\alpha_i^*\left\langle x,x_i\right\rangle+b^* \tag{28}$$

The optimization technique which is given above is known as Maximal Margin Classifier (MMC). In SMC, however, tolerances are accepted and the separating hyperplane is constructed with some degree of error. For SMC setting, the primal form of the optimization problem can be written as,

$$\min_{w,b}\frac{1}{2}\|w\|^2+C\sum_i\xi_i \quad st: y_i\left(\left\langle w,x_i\right\rangle+b\right)\geq 1-\xi_i,\ \xi_i>0,\ i=1,..,m \tag{29}$$

where $\xi_i$ denotes the slack variable, $C$ is regularization parameter which is used to control the training error and model complexity. Dual form of the problem is same as MMC,

$$\max_{w,b}\sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i,j}^{m}\alpha_i\alpha_j y_i y_j\left\langle x_i,x_j\right\rangle \tag{30}$$

where $\alpha$ is the vector of Lagrange multiplier and $m$ is number of input data vectors subject to the conditions in (31).

$$0\leq\alpha_i\leq C, i=1,\ldots,m\ \ and\ \sum_{i=1}^{N}\alpha_i y_i=0 \tag{31}$$

The hyperplane decision function is the same as that in (28). As we can see in (27) and (30), separating plane is described by $\alpha$ values. Support vectors can be described as data the associated Lagrange multipliers of which are positive.

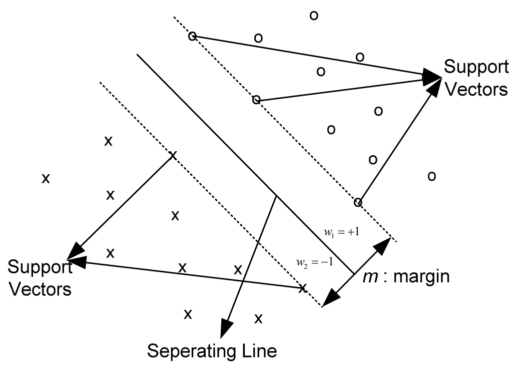


Fig. 6: 2-D linearly separable case

In figure 4, there is a boundary near the separating line, and data which are lying on the boundary line are support vectors.

### *4.2.2 Nonlinear Classification with SVM*

In practical applications, the classes available implicitly in the data are not linearly separable and in such cases linear classifiers perform poorly. Nonlinear classifiers in these circumstances are useful as exemplified by Boser et al

(1992) and Chen et al (2005). Nonlinear kernel functions are used for mapping the input data to a high dimensional Euclidean space where a possible solution will be sought. Two of the frequently used kernel functions are given in (32) and (33),

$$K\left(\mathbf{x}_i,\mathbf{x}_j\right)=\exp\left(\frac{-\left\|\mathbf{x}_i-\mathbf{x}_j\right\|^2}{2\sigma^2}\right) \tag{32}$$

$$K\left(\mathbf{x}_i,\mathbf{x}_j\right)=\left(\mathbf{x}_i^T\mathbf{x}_j+1\right)^p \tag{33}$$

Primal form of the optimization problem is the same as linear classifier; the change in the expression is that kernel function is used for inner product, and the dual form of the optimization problem will be as given in (34),

$$\max_{w,b}\ \sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i,j}^{m}\alpha_i\alpha_j y_i y_j k\left(x_i,x_j\right) \tag{34}$$

where $k(x_i,x_j)$ is the kernel function and the conditions are the same as linear case (See 31).

*4.2.2.1 υ–Support Vector Classifier*

υ Support Vector Classifier (υ - SVC) is a modified version of nonlinear SMC. The regularization parameter $C$, is replaced by a parameter $\upsilon\in[0,1]$. $C$ parameter in SMC is used to punish the error to decrease, but υ parameter is used to determine upper bound of the error and lower bound of the support vectors which are used to create separating hyper plane, so in υ–SVC, we can determine the number of support vectors. Primal form of the optimization problem is changed to given form as following,

$$\begin{aligned}&\min_{w,\xi,\rho,b}\frac{1}{2}\|w\|^2-\upsilon\rho+\frac{1}{m}\sum_{i=1}^{m}\xi_i\\&st: yi(\langle w,xi\rangle+b)\geq\rho-\xi_i,\ i=1,\ldots,m\\&\quad\xi_i\geq 0,\ \rho\geq 0\end{aligned} \tag{35}$$

where $\rho$ and $\upsilon$ are optimization parameters and the dual form changes to,

$$\begin{aligned}&\max_{\alpha\in\mathbb{R}^m}\ -\frac{1}{2}\sum_{i,j=1}^{m}\alpha_i\alpha_j y_i y_j k(x_i,x_j)\\&st:\quad 0\leq\alpha_i\leq\frac{1}{m},\qquad\sum_{i=1}^{m}\alpha_i y_i=0,\qquad\sum_{i=1}^{m}\alpha_i\geq\upsilon\end{aligned} \tag{36}$$

where $\alpha_i$s are Lagrangian multipliers which are calculated by solving (36), and the decision function is as given below.

$$f(x)=\operatorname{sgn}\left(\sum_{i=1}^{m}\alpha_i y_i k\left(x,x_i\right)+b\right) \tag{37}$$

*4.2.2.2 Least Squares Support Vector Classifier*

LS-SVC is more popular than Vapnik's SVM classifier, because in the latter, Quadratic Program (QP) solvers are needed and QP solvers introduce significant computational intensity. Because of this, LSSVC, proposed by Suykens (2002), uses simple matrix calculations to obtain the solution of the problem. Suykens (2002) modifies the original optimization as given in (38),

$$\begin{aligned}&\min_{w,b,e} J_P(w,e)=\frac{1}{2}w^T w+\gamma\frac{1}{2}\sum_{k=1}^{N}e_k^2\\&st\quad: y_k[w^T\varphi(x_k)+b]=1-e_k,\ k=1,\ldots,N\end{aligned} \tag{38}$$

where $N$ is the number of data pairs, $e_k$ is the associated error value, $\gamma$ is a parameter establishing a balance between the complexity and the performance. If dual form of the optimization problem is calculated using Lagrange multipliers one gets the matrix equality in (39),

$$\begin{bmatrix}0 & y^T\\ y & \Omega+I/\gamma\end{bmatrix}\begin{bmatrix}b\\ \alpha\end{bmatrix}=\begin{bmatrix}0\\ 1_v\end{bmatrix} \tag{39}$$

where $y$ is the vector of target values (the desired class label), $I$ is the identity matrix of appropriate dimensions, $\gamma$ is LS-SVC parameter, $b$ is the bias, $1_\upsilon=[1,\ldots,1]$, $k(x_k,x_l)$ is the kernel function and $\Omega_{kl}:=y_k y_l k(x_k,x_l)$ is obtained. Decision function is given by (40),

$$y(x)=\operatorname{sgn}\left(\sum_{i=1}^{N}\alpha_k y_k k(x,x_i)+b\right) \tag{40}$$

## 5. EXPERIMENTS

In this section, the presented classifiers are examined with features extracted from a video sequence. During the training of the all classifiers, a set of 160 data pairs is used for training and 61 data is used to test the system performance. 23 of the training data include the features of the objects which are not car, and 38 of the training data include the features of the objects which are car. The data may be seen unreliable in number; however, distribution of the data on illumination, rotation, and translation changes is quite well to obtain reliable points on feature space for implementing intelligent systems successfully. The features of segmented masks are used to be able to test the trained system. Fig. 3 gives some examples of tested segmented masks obtained from the video sequences. The output is +1 if the mask contains a car; otherwise the output is −1 indicating that the object is not a car. The test results with misclassifications are tabulated in Table 1.

**Table 1** A Comparison of the Classification Performances

| | **#Correct Classifications** | **#Misclass.** | **Percent Accuracy** |
|---|---|---|---|
| ANN | 54 | 7 | 88.53% |
| υ-SVC | 57 | 4 | 93.44% |
| LS-SVC | 60 | 1 | 98.36% |

According to the results, all three systems misclassify sometimes. One prime reason for this is the inappropriate features due to the illumination change during the capturing of frames. Also, segmentation of the RoI has an effect on feature extraction. The area of the segmented region may also result in misclassification. Another reason is the noise in the segmented RoI caused by illumination changes and weakness of the camera system. Although there is some false recognition, all three systems have a reliable recognition rate. ANN is the least accurate system among the three methods due to a misclassification rate of 11.47%.

Extracting better features and changing the activation function may lead to some increase the accuracy of ANN having 8 hidden neurons in the hidden layer. According to the results, SVM algorithms give a better recognition rate with respect to ANN structure. The best accuracy is obtained with LS-SVC with an error rate of 1.64%. However, υ-SVC has less computational complexity than LS-SVC. Similarly, the performances of both SVCs might be enhanced by changing the kernel function and its kernel parameters.

## 6. CONCLUSIONS

This work presents the comparison of three methods to classify objects within the frame of a video sequence. The question was to seek for a car in the RoI and ANNs, LS-SVC and υ-SVC are tested for several experimental data. It is seen that, though practical, ANNs perform poorly in terms of the SVCs, among which best performance is obtained with LS-SVC and least complexity is obtained with υ-SVC.

The authors aim at implementing these schemes on a quadrotor type unmanned aerial vehicle to be used for autonomous object tracking.

## 7. ACKNOWLEDGMENTS

This work is supported by TÜBİTAK 1001 programme, grant no 107E137. The authors gratefully acknowledge the facilities of the Unmanned Air Vehicles Laboratory of TOBB Economics and Technology University.

## REFERENCES

Andrade, M.C., (2004). An Interactive Algorithm for Image Smoothing and Segmentation. *Electronic Letters on Computer Vision and Image Analysis*, **4**(1), 32-48.

Borchani, M. and Stamon, G. (1997). Use of Texture Features for Image Classification and Retrieval. *In Proceedings of SPIE Conference on Storage and Retrieval for Image and Video Databases III*, **3229**, 401-406.

Burges, J.C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, **2**, 121-167.

Boser, B.E., Guyon, I.M. and Vapnik, V. (1992) A training algorithm for optimal margin classifiers. *In Fifth Annual Workshop on Computational Learning Theory*, Pittsburgh, ACM.

Chen, P., Lin, C., Schölkopf, B. (2005). A Tutorial on υ - Support Vector Machines. *Applied Stochastic Models in Business and Industry*, **21**(2), 111-136.

Donoho, D.L., (1995). De-Noising by Soft Thresholding. *IEEE Transactions on Information Theory*, **41**(3), 613-627.

Elgammal, A., Duraiswami, R., Harwood, D., Davis, L.S. (2002). Background and Foreground Modelling Using Nonparametric Kernel Density Estimation for Visual Surveillance. *Proceedings of the IEEE*, **9(**7), 1151-1163.

Gonzalez, Rafael, and Woods, R.E. (2008). *Digital Image Processing*, Pearson Prentice Hall, Upper Saddle River, New Jersey.

Gould, S., Arfvidsson, J., Kaehler, A., Sapp, B., Messner, M., Bradski, G., Baumstarck, P., Chung, S., and Ng, A.Y. (2007). Peripheral-Foveal Vision for Real-time Object Recognition and Tracking in Video. *IJCAI*, 2115-2112.

Hagan, M.T. and Menhaj M.B. (1994). Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, **5**(6), 989-993.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*, Pearson, Prentice Hall, Upper Saddle River, New Jersey.

Irani, M. and Peleg, S. (1993). Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency. *Journal of Visual Communication and Image Representation*, **4**(4), 324-335.

Jamarani, S.M.H., Moradi, M.H., Behnam, H. and Rezai Rad, G.A. (2005). Intelligent System for Breast Cancer Prognosis using Multiwavelet Packets and Neural Network. *Proceedings of World Academy of Science, Engineering and Technology*, **7**, 451-456.

Lourakis, M.I.A., Argyros, A.A. (2005). Is Levenberg-Marquardt the Most Efficient Optimization Algorithm for Implementing Bundle Adjustment?. *Proceedings of the Tenth IEEE International Conference on Computer Vision*, **2**, 1526-1531.

Niu, W., Jiao, L., Han, D., and Wang, Y. (2003). Real-Time Multi-person Tracking in Video Surveillance. *ICICS-PCM*, **2**, 1144-1148.

Nixon, M. and Aguado, A. S. (2008). *Feature Extraction and Image Processing*, Elsevier Ltd., Theobald's Road, London.

Piccardi, M. (2004). Background Subtraction Techniques: A review. *IEEE International Conference on Systems, Man and Cybernetics*, **4**, 3099-3104.

Russel, S., Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*, Pearson, Prentice Hall, Upper Saddle River, New Jersey.

Sarlashkar, M. N. Bodruzzaman, M., and Malkani, M.J. (1998). Feature Extraction Using Wavelet Transform For Neural Network Based Image Classification. *Proceedings of the Thirteenth Southeastern Symposium on System Theory*, 412-416.

Suykens, J.A.K., Gestel, T.G., De Brabanter, J., De Moor, B., Vandewalle, J. (2002). *Least Squares Support Vector Machines*, World Scientific Publishing Co. Pte. Ltd., Singapore.

Töreyin, B.U. (2003). *Moving Object Detection and Tracking in Wavelet Compressed Video*. Graduation Thesis of Department of Electrical and Electronics Engineering, Bilkent University.

Walker,J.S. (1999). *A Primer on Wavelets and their Scientific Applications*, Chapman & Hall/CRC, Boca Raton.

Zhang, G. P. (2000). Neural Networks for Classification: A Survey. *IEEE Transactions on Systems, Man, And Cybernetics, Applications And Reviews*, **30**(4), 451-462.

Zivkovic, Z. (2004). Improved Adaptive Gaussian Mixture Model for Background Subtraction. *In Proceedings: ICPR*, **2**, 28-31.