# Experiments on Individual Classifiers and on Fusion of a Set of Classifiers

Claude Tremblay [1,2]
[1]Centre de Recherches Mathématiques
Université de Montréal
C.P. 6128 Succ. Centre-Ville,
Montréal, QC, H3C 3J7, CANADA
claude.tremblay@umontreal.ca

Pierre Valin [2]
[2]Lockheed Martin Canada
6111 Royalmount avenue
Montréal, QC, H4P 1K6, CANADA
pierre.valin@lmco.com

**Abstract** – *In the last decades many classification methods and fusers have been developed. Considerable gains have been achieved in the classification performance by fusing and combining different classifiers. We experiment a new method for ship infrared imagery recognition based on the fusion of individual results in order to obtain a more reliable decision [1]. To optimize the results of every class of ship, we implemented individual classifiers using Dempster-Shafer(DS) method for each class i.e. an individual classifier returns if the ship belongs to the class or not. We compare the result of the DS classifier with the results of the individual classifier. The improvement recognition varies between 3% to 20% for a class. We then experiment a new method based on a fusion of a set of classifiers [2]. The objective of a good fuser is to perform at least as good as the best classifier in any situation. For this purpose, we consider three classifiers: DS classifier, Bayes classifier and nearest neighbor classifier and one fuser: feedforward neural network fuser. We compare the results of the best classifier with the results of the fusion of a combination of classifiers. The fuser gives a performance equal or superior to the best classifier.*

**Keywords:** FLIR imagery, fusion of classifiers, Bayes classifier, DS classifier, k-nearest neighbors classifier, neural networks fuser.

## 1 Introduction

In image recognition, several classifiers have been designed and implemented and their performance is varied. A solution to various performances is to fuse or combine different classifiers in order to get better performance compared to the best classifier.

Recently, Rao has demonstrated that individual results can be fused in order to obtain a more reliable decision and he has also demonstrated that the performance of a fuser can be guaranteed to perform at least as good as the best classifier under certain conditions [3]. Our work is closely related to Rao's work but in a more practical way. First, we compare the result of the Dempster-Shafer (DS) classifier with the result of a fusion of individual DS classifiers. We then consider three classifiers: DS classifier, Bayes classifier and nearest neighbor classifier. We compare the results of the classifiers with the results of the fusion of two or three classifiers with a neural network fuser.

## 2 Data Set and Feature Selection

Park and Sklansky [4] have developed an automated design of linear tree classifiers for ship recognition. We used the same data set and features in our work.

The data set is composed of 2545 forward looking infrared (FLIR) ship images. Each ship image belongs to one of the eight classes listed in Table 1. For every image, the ship silhouette was threshold manually. Figure 1 shows silhouettes for the 8 classes of Table 1.

The features used are seven invariant moments given by Hu [5]. These moments are invariant under translation, rotation and scale. But these moments deliver information primarily of the global shape of the object and represent poorly the details of the objet. Hence, Park and Sklansky added as features four parameters extracted by fitting an auto regressive model to one-dimensional sequence of the projected image along the horizontal axis.

| Class | Class of ship |
|---|---|
| 1 | Destroyer |
| 2 | Container |
| 3 | Civilian Freighter |
| 4 | Auxiliary Oil Replenishment |
| 5 | Landing Assault Tanker |
| 6 | Frigate |
| 7 | Cruiser |
| 8 | Destroyer with Guided Missile |

Table 1: Ship classes

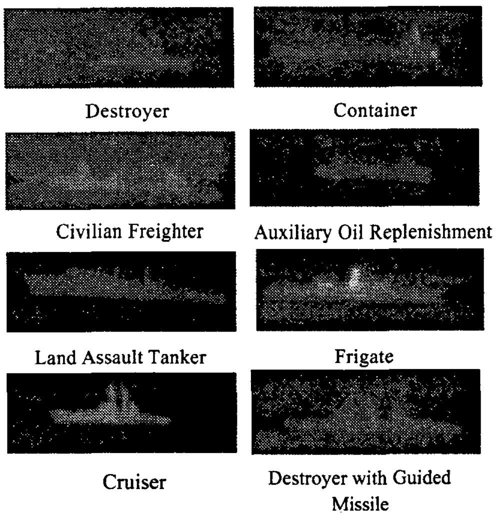


Figure 1: Images of the 8 classes of ships

The quality of each image varies a lot. The distance of the ship to the camera and the noise on the image affects the performance of the classifier.

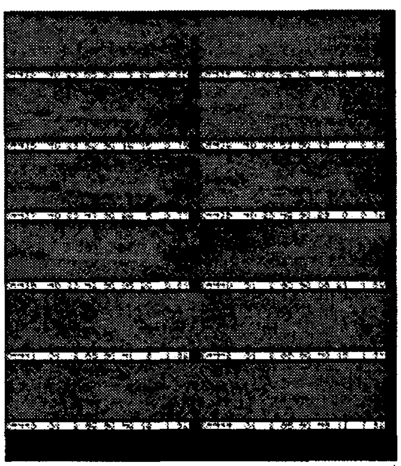

Figure 2: Civilian Freighter in broadside view

# 3 Classifiers and fuser description

## 3.1 Bayes Classifier

Bayes classifiers use a probabilistic approach to assign a class. They compute the conditional probabilities of different classes given the values of the attributes and then predict the class with the highest conditional probability.

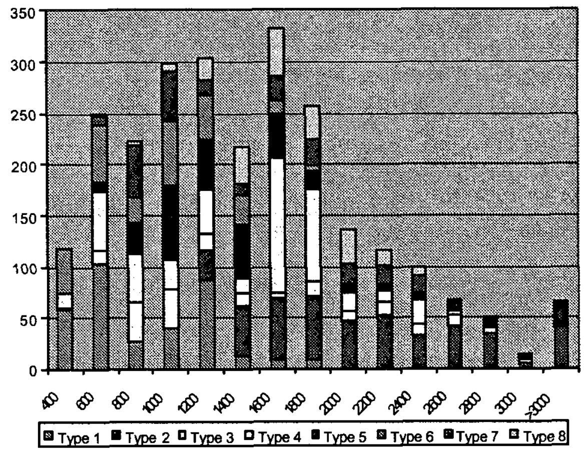


Figure 3 : Frequency graph for attribute 1

Equation (1) represents the probability of an object belonging to in the class i ($C_i$) knowing the value of attribute j ($A_j$), where i represents the number of classes i = {1, 2, ...,m} and j the number of attributes = {1,2,..., N}.

$$P(C_i|A_j) = \frac{P(A_j|C_i)\,P(C_i)}{\sum_{i=1}^{m} P(A_j|C_i)P(C_i)} \quad (1)$$

We compute the probability of an object to be in the class i knowing the value of attribute j for each attribute and sum them

$$P_s(C_i)=\sum_{j=1}^{N} P(C_i|A_j) \quad (2)$$

Finally, we identify the class of the object X. We choose the class with the highest probability.

$$X=\mathrm{Arg}\{\max_{1\le i\le m} [P_s(Ci)]\} \quad (3)$$

## 3.2 k-Nearest Neighbors Classifier

The k-nearest neighbor classifier finds the k nearest neighbors based on a metric distance and returns the class with the greatest frequency.

We used a distance weighted by the inverse of the inter-classes covariance matrix:

$$d_{\Gamma}^{2}(x_1, x_2) = (x_1 - x_2)^{T}\,\Gamma^{-1}(x_1 - x_2) \qquad (4)$$

where the inter-classes covariance matrix is defined by:

$$\Gamma = \frac{\sum_{i=1}^{N}(x_i - \overline{x})^{T}(x_i - \overline{x})}{N-1} \qquad (5)$$

and where $x_1, \ldots, x_N$ are N vectors, for which we know the true identity.

## 3.3 DS Classifier

DS theory of evidence is a good means of reasoning under uncertainty. A key aspect of this theory is its ability to combine evidences by using the technique of orthogonal summation.

The DS theory requires that the propositions pertain to the set of all possible propositions that can be output. This set is called the frame of discernment denoted by $\theta$ whose elements are mutually exclusive. The power set of $\theta$ is $P(\theta)$. $P(\theta)$ is the set of all the $2^{\theta}$ subsets of $\theta$. Let A be an element of $P(\theta)$.

A basic probability assignment is a function from $P(\theta)$ to [0, 1] is defined by :

$$m:\ P(?) \rightarrow [0, 1]$$

$$A \rightarrow m(A) \qquad (6)$$

There are three fundamental axioms about the mass:

$$m(\phi) = 0 \qquad (7)$$

$$\sum_{i=1}^{2^{|\theta|}} m(A_i) = 1 \qquad (8)$$

$$m(A_1 \cup \ldots \cup A_n) \geq \sum_{I \subset \{1,2,\ldots,n\}} (-1)^{|I|+1}\, m\Big(\bigcap_{i \in I} A_i\Big) \qquad (9)$$

The probability distribution of $P(\theta)$ can be estimated by the mass function. The precise probability distribution of $P(\theta)$ may not be known exactly so bounds of probability distribution are defined. The lower probability and upper probability of a subset A of $P(\theta)$ is denoted as belief measure Bel(A) and plausibility measure Pls(A), respectively. They can be determined from the mass function as follows:

$$Bel(A) = \sum_{i=1}^{2^{|A|}} m(A_i) \qquad (10)$$

$$Pls(A) = 1 - Bel(\neg A) \qquad (11)$$

Generally, $Bel(A) \neq Pls(A)$, the true probability of A is between Bel(A) and Pls(A).

The combined mass functions of two independent mass function $m_1$ and $m_2$ is calculated by using DS's rule of combination, denoted by $m_1 \oplus m_2$:

$$m_1 \oplus m_2(A) = \frac{\sum_{B \cap C = A} m_1(B)\, m_2(C)}{1-K} \qquad (12)$$

$$K = \sum_{B \cap C = \varnothing} m_1(B)\, m_2(C) \qquad (13)$$

where K is normalization constant, called conflict because it measures the degree of conflict between B and C. After the combination, we return the class with the highest mass.

## 3.4 Single fuser

The performance of fusion of classifiers has been demonstrated by analytical and experimental results. The choice of a fuser is very important because a bad fuser choice can result in a performance worse than the worst classifier. Figure 4 shows the concept of the fusion of the classifiers.

We choose the neural network fuser because this fuser is guaranteed to give results at least as good as the best classifier.

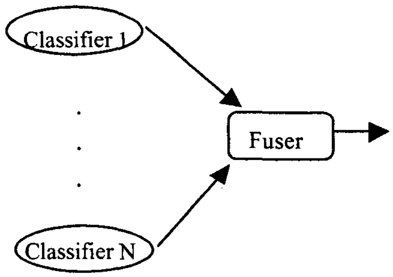


Figure 4: Single fuser

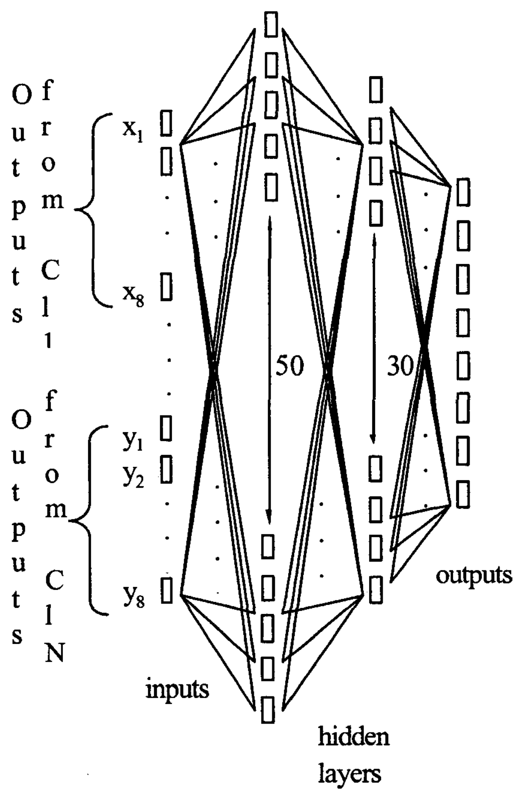


Figure 5: Neural network structure

where Cl i means classifier i.

# 4 Experiments on individual classifiers

## 4.1 Results of the DS Classifier

We implemented a classifier using the DS method. We fused sequentially the eleven attributes with this method. Table 2 shows the confusion matrix:

| | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 |
|---|---|---|---|---|---|---|---|---|
| C1 | 0.87 | 0.01 | 0.00 | 0.01 | 0.00 | 0.08 | 0.03 | 0.00 |
| C2 | 0.00 | 0.86 | 0.00 | 0.10 | 0.04 | 0.00 | 0.00 | 0.00 |
| C3 | 0.04 | 0.08 | 0.07 | 0.64 | 0.17 | 0.00 | 0.00 | 0.00 |
| C4 | 0.00 | 0.04 | 0.00 | 0.96 | 0.00 | 0.00 | 0.00 | 0.00 |
| C5 | 0.00 | 0.20 | 0.00 | 0.21 | 0.63 | 0.00 | 0.00 | 0.00 |
| C6 | 0.15 | 0.01 | 0.00 | 0.01 | 0.05 | 0.74 | 0.00 | 0.05 |
| C7 | 0.33 | 0.11 | 0.00 | 0.01 | 0.05 | 0.00 | 0.49 | 0.01 |
| C8 | 0.01 | 0.01 | 0.00 | 0.00 | 0.01 | 0.09 | 0.00 | 0.88 |

Table 2: DS Classifier confusion matrix

## 4.2 Results of Individual DS Classifiers

To improve the results of every class of ship, we implemented individual classifier using the DS method for each class i.e. an individual classifier returns if the ship belongs to the class X or not. For every individual classifier, we chose a subset of features, which optimize the performance of the class. Table 3 shows the percentage of correct classification of the DS method, of each individual DS. We see that individual DS classifier gives better results for all the classes than the DS classifier.

| | DS Classifier | Individual DS class. |
|---|---|---|
| Class 1 | 0.871 | 0.932 |
| Class 2 | 0.864 | 0.958 |
| Class 3 | 0.070 | 0.242 |
| Class 4 | 0.957 | 0.984 |
| Class 5 | 0.629 | 0.747 |
| Class 6 | 0.735 | 0.803 |
| Class 7 | 0.490 | 0.686 |
| Class 8 | 0.885 | 0.928 |
| Total | 0.745 | ---- |

Table 3: Classification (with DS) and fusion results

# 5 Experiments on a neural network fuser for a set of classifiers

## 5.1 Results of Single Classifiers

First, we tested the data set with three methods of classification: Bayes, DS and knearest neighbor. The classification results of each method are listed in table 4. For the nearest neighbor method, we used k = 3 and the weighed distance by the inverse of the inter-classes covariance matrix.

| Classification method | 1000 images | 1500 images | Total |
|---|---|---|---|
| Bayes | 0.777 | 0.773 | 0.745 |
| DS | 0.745 | 0.746 | 0.774 |
| 3-Nearest neighbor | 0.946 | 0.950 | 0.948 |

Table 4: Classification results

## 5.2 Results of the fusion of classifiers

We then fused the results of two or three classification methods with a feedforward neural network fuser. Our neural network fuser has 16 or 24 inputs (these inputs are the results of selected subsets of two or three classifiers) and has 8 outputs, one for each class. We used the following parameters: 2 hidden layers, 50 neurons on the first layer, 30 neurons on the second layer momentum = 0.5, maximal error = 0.001, epsilon = 0.1, number of maximal iterations = 100. We trained the fuser with 1000 data and tested on 1500 data. We also trained the fuser with 1500 data and tested on 1000 data. From tables 5 and 6, we can see that the fuser gives performance equal or superior than the best classifier.

| Training size | Testing size | Bayes, DS | Bayes, K-NN | Bayes , K-NN, DS |
|---|---|---|---|---|
| 1000 | 1500 | 0.814 | 0.950 | 0.951 |
| Best single classifier | | 0.773 | 0.950 | 0.950 |

Table 5: Fusion results of classifiers with feedforward neural networks.

| Training size | Testing size | Bayes, DS | Bayes, K-NN | Bayes, K-NN, DS |
|---|---|---|---|---|
| 1500 | 1000 | 0.853 | 0.955 | 0.956 |
| Best single classifier | | 0.777 | 0.946 | 0.946 |

Table 6: Fusion results of classifiers with feedforward neural networks.

# 6 Conclusions

The results indicate that individual classifiers can be a good choice. In our particular case, the individuals DS classifiers perform better. An advantage of this method is that we use simple algorithms.

We see that a feedforward neural network is a good choice for a fuser. In our experiments, the performance of the fuser was always at least as good as the best classifier. Fusion of classifiers is a promising technique for image recognition.

# 7 Acknowledgements

FLIR images from the United States Naval Airfare Center, China Lake, California was provided by Dr. Jack Sklansky of the University of California at Irvine.

# 8 References

[1] J. Kittler and F. Roli. *Multiple Classifiers Systems*, volume 1857. Springer-Verlag, Berlin, 2000.

[2] N. S. V. Rao. *On Sample-Based Implementation of Decision Fusion Functions*, Workshop on Target Tracking and Sensor Fusion, Monterey, CA, May 2001.

[3] N. S. V. Rao. *On Design and Performance of Metafusers*, Proceedings of the Workshop on Estimation, Tracking and Fusion: A tribute to Yaakov Bar-Shalom, Monterey, CA, May 2001.

[4] Y. Park and J. Sklansky. *Automated Design of Linear Tree Classifiers*, Pattern Recognition, Vol. 23, No. 12, pp.1393-1412, 1990.

[5] M. K. Hu. *Visual Pattern Recognition by moment invariant*, IRE Trans. Inform. Theory, IT-8, pp.179-187, 1962.

[6] F. Rhéaume, A-L Jousselme, D. Grenier, E. Bossé, and P. Valin, "New Initial Basic Probability Assignments for Multiple Classifiers", in Conference 4729, *Signal Processing, Sensor Fusion, and Target Recognition XI*, SPIE Aerosense 2002, Orlando, USA, April 1-5 2002, paper #35.