The Complexity of the Algorithms for the Image Recognition and Classification (IRC)

IOAN ISPAS¹, EDUARD FRANTI², FLORIN LAZO²

¹Petru Maior University, N.Iorga 1, Tirgu Mures, ROMANIA,
²Research Institute for Artificial Intelligence, 13 Septembrie, Bucharest, ROMANIA,

Abstract: - The problem of the image recognition and classification (IRC) based on the pattern recognition is of a strategic importance in lots of domains. Our contribution consists in the foundation of the complexity of the algorithms for image recognition and classification. This will help us issue some precise statements on the computational difficulty of the problem of the image recognition and classification (IRC).

Key-Words: - modelling, image recognition and classification, algorithm complexity

1 Introduction

A short analysis of the problem of image recognition and classification (IRC)

The automatic classification of the images is of a strategic importance in lots of domains. Its solving is based on the methods and algorithms of automatic pattern/object recognition and image classification [3].

In the following part, we will define the IRC problem in an abstract form which contains the essence of the real world problem:

*Given a image data base (data stream) \( B = \{I_1, I_2, ..., I_n\} \) containing a 'main character'; given a set of descriptions of some known distinct objects \( R = \{O_1, O_2, ..., O_k\} \); knowing that any human operator is able to recognized easily, by means of rapid visual inspection, the objects in the image. The aim of the algorithm is to determine the images that contain these objects and to classify them in \( k+1 \) distinct classes: \( C_1, C_2, ..., C_k \) and \( C_{k+1} \). The classes \( C_i \), \( i=1,k \) will group all the images containing the objects \( O_i \), \( i=1,k \), and the class \( C_{k+1} \) will group the images without any of the \( R \) objects. The diagram of the IRC problem is the following (figure 1):

![Fig.1 Diagram of the IRC problem](image)

The next table offers a small example consisting of eight images with 'characters' [5] that can be classified into seven image classes: 1. the class of the images containing horses; 2. the class of the images containing cheetahs; 3. the class of the images containing elephants; 4. the class of the images containing airplanes; 5. the class of the images containing bears; 6. the class of the images containing eagles; 7. the class of the images „neutral”, without recognized object. In this case, the component elements of the set of the recognized 'objects' are: 1. horses; 2. cheetahs; 3. elephants; 4. airplanes; 5. bears; 6. eagles (figure 2).

![Fig.2 Images of 'objects'](image)
Having studied the specialized literature [1, 2] we can state that the image recognition algorithms describe a four-stepped process [3]. Each step is essential and inevitable. The diagram of the image recognition algorithms is the following (figure 3):

Fig.3 Diagram of the image recognition algorithms

A. Preprocessing of the image. This means the application of some DIP algorithms specialized in enhancing image quality [1].

B. Feature extraction. This is the key step, the one that measures the performances and the quality of the recognition software. The discovery of the most proper features and characteristics is the very key of the success [4, 6, 7, 8, 9, 10, 11, 12]. The final result of this step is a feature vector description \( (v_1, v_2, ..., v_n) \), not necessarily numerical.

C. Feature/pattern measurement. This step is well theoretically founded; there is a developed mathematical theory (The measure theory) which can help us select the proper and efficient n-dimensional metrics. The final result of this step is usually a one- or multi-dimensional value (a vector) perceived as the ‘distance’ of the feature vector towards the borders of the class [1, 4, 7, 10, 11].

D. Image/Pattern classification. This is the final step which combines the results of the prior measurements. This establishes the fact that the pattern/object – described by the feature vector – belongs to distinct class of images, according to certain appartenance mathematical criteria. The final result of the classification step is number C of the image class.

2 The simplified version of the problem of the image recognition and classification (sIRC)

If we consider the objects-‘characters’ from the images as marks/signatures, then the recognition of images leads us to the following simplified version of the problem:

Given a image data base (or a image stream) \( B = \{I_1, I_2, ..., I_n\} \) containing a ‘main character’ that marks them; given a set of descriptions of some known distinct objects \( R = \{O_1, O_2, ..., O_k\} \). The following algorithm determines the images from B that contain these objects and classifies them in \( k+1 \) distinct classes: \( C_1, C_2, ..., C_k, C_{k+1} \). The class \( C_i \) will group all the images containing the object \( O_i \), \( i=1..k \) and the class \( C_{k+1} \) will group the images without any of the \( R \) objects.

Algorithm sIRC(image I);
For \( i=1..k \) do
  If Recognition\( (O_i, I) \) return\( (C_i) \);
Return\( (C_{k+1}) \);

The equivalence between the IRC problem and its simplified version \( \text{sIRC} \) still remains an open question. Our belief, just like its title shows, is that the simplified version \( \text{sIRC} \) problem is less difficult than the initial one. Unfortunately, we cannot proof rigorously this statement although the multitude of facts strongly confirms it.

We will later on focus on the study of the complexity of the \( \text{Classification_sIRC} \) algorithm. It is obvious that its complexity relies on the complexity of the \( \text{Recognition} \) sub-algorithm. The total complexity of the algorithm is in the worst case:

\[
\text{WorstCase(Classification_sIRC)} = k \times O(\text{Recognition})
\]

The \( \text{Recognition} \) algorithm is the clue of the \( \text{sIRC} \) problem. Its input is I image and the description of the recognition pattern/object O. Its output will be \text{true} or \text{false}.

Considering that any human operator is able to easily recognize the presence of the object O in the image I by visual inspection, based on a primary process of the \text{mathematical modeling and formalization of the sIRC problem}, we can state the following.

2.1. Statements resulting from the mathematical modeling of the \( \text{sIRC} \) problem

S1. The object \( O_k \) ‘marks/signs’ singularly the image \( I_k \);  

S2. The object \( O_k \) has its own identity, which makes it distinctly distinguishable;  

S3. The identity of a certain object \( O \) is given by two
independent features: its unique pattern \( P \) and its information content (color) \( C \).

While the \( P \) pattern distinctly delimits it from the environment, its information content (colorist) \( C \) identifies it: \( O = O ( P, C ) \).

For instance, the next image in figure 4, contains the ‘character’ horse. We deliberately introduced its shadow below, in order to relief the shape of the ‘object’. Still, only the information content – colorist \( C \) of the ‘object’ - makes it distinct from its shadow!

\[
\text{Fig.4 ‘Chararcter’ horse}
\]

S4. The shape/pattern and the content of a certain object \( O \) is described by a pair of independent ‘codes’/descriptors \( (P, C) \); 

S5. each object \( O \) is uniquely correspondent to a point \( O(P,C) \) in the bidimensional searching space given by the ordonates \( P \) and \( C \), one for patterns (forms) and the other for information contents (colouristics).

Important notice. These statements have a mandatory character and they are independent one another.

2.2. Definitions resulting from the mathematical modeling of the sIRC problem

D1. By the searching space we understand a set of data \( S \) which has to be exhaustively covered in order to find the target data \( x \) among \( S \) data. 

Given \( n \) the cardinal of the set \( S \). Considering that the exhaustive covering condition of the searching is needed, the number of required steps (comparisons) in order to find out \( x \) is in the worst case \( n \).

D2. By the pattern \( P \) of an (bi-dimensional) object we understand the set of the contour points (laying on the external edges of the shape of the object) which delimits the space occupied by it. The pattern \( P \) of an object is that what makes it distinguished from the environment and confers its identity.

D3. By its information content (colorist) \( C \) of a certain object we will understand a set of points belonging to the object, grouped together according to an association (relational) criteria. 

For instance, the set of the ‘interior’ points of the object, the set of the points of the same color, etc. The information content (colorist) \( C \) is the visible, descriptive expression of the object.

2.3. Conclusions resulting from the mathematical modeling of the sIRC problem.

C1. Generally speaking, for each Recognition\((O,I)\) algorithm the image \( I \) generates a searching space for the object to be recognized \( O(P,C) \).

Proof. Considering that the object is described by the two independent vectors within the co-ordinate space \( P-C \), we can approach it as a rectangle. The recognition of the object \( O \) into \( I \) then becomes equivalent with the recognition of a rectangle \( R \) having the color \( C \). By \( D \) we might understand either the ‘description’ of the shape of the rectangle, or its dimension – length/width.

Any algorithm, according to the Turing-Church thesis [13], has to be considered as a Turing machine having as an input the sequential description of the image \( I \) and of the object \( O(P,C) \), respectively the rectangle \( (R,C) \). Thus, it is obvious that any algorithm Recognition\((R, C)\) cannot leave any sub-image (sequence) of the image \( I \) of the at least \( D \) dimension un-inspected. This sub-image could contain the rectangle \( (R, C) \). Therefore, the exhaustive scanning of the image becomes a genuine necessity!

This leads us to the obvious conclusion that the image \( I \) becomes a searching space for the object to be recognized \( O(P,C) \).

C2. Features recognition, pattern and color, described by the pair of codes \( (P, C) \) in the searching space of the image \( I \), is equivalent with the seeking of a target text in a source text page. The approximate matching rank has to be more then a certain threshold.

Proof. Because the object to be recognized is described by its two features - the pair of descriptors (codes) \( (P, C) \) – according to (C1) the recognition process consist of an exhaustive scanning and an approximate matching of the codes \( (P, C) \) in the searching space of the image \( I \). This process is
identical with the searching of an approximate matching of a target text within a source text.

C3. The complexity of the algorithm
Recognition(O, I) is direct proportional with the dimension of the image I and with the dimension of the codes P and C:

O( Recognition(O,I) ) = O( Dim(I) x Dim(P) x Dim(C) )

3 Conclusion

R1. The difficulty of the simplified version of the problem siRC is a consequence of the complexity of the algorithm Recognition(O, I).

R2. The essential step in the recognition process of the object O is the feature extraction of the patterns and of the information content (P, C) from the image I.

Notice. This feature extraction step is inevitable because the image I is formed by a matrix of pixels, but the descriptors (P, C) are a pair of codes describing the shape and the information content of the object O.

R3. The complexity of the algorithm Recognition(O,I) is directly proportional with the dimension of the features (P, C) extracted by the sub-algorithm FeatureExtraction(O).

R4. The complexity of the algorithm Recognition(O,I) is given by the formula:
O(Recognition(O,I)) = O( Dim(I) x O(FeatureExtraction(O)) )

Important notice. The extraction of the color and of the pattern features from the image may imply a very consistent number of operations m (i.e. associations and relations) over the pixels within the interest zones. The complexity of the extraction algorithm becomes:

O(FeatureExtraction (O)) = m x Dim(ExtractionZone)

Note that the determination/discrimination of the interest zones (which could contain the object) is the most important but also the most difficult step in the entire feature extraction process.

This can lead to a situation wherein the recognition of an object having dimension 200 x 200 pixels, within an image having a resolution of 800 x 600 pixels and 256 colors, could require a number of operations directly proportional with the huge value 800 x 600 x 256 x 200 x 200, greater than 10^{10}.

The final conclusion about the difficulty of the image recognition and classification problem is that a proper solution of the problem and of its simplified version siRC depends in the most direct way on the design of an efficient pattern/information content extraction sub-algorithm.

References:
