IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 14, Issue 1 Ver. I (Jan. - Feb. 2018), PP 65-71
www.iosrjournals.org

Application of rough sets theory in image classification

Ahmad Abd Allah Mohamed Zedan

Department of Mathematics, College of Science, Majmaah University, Majmaah 11952, Saudi Arabia

Abstract: *In this work we study the relationship between rough sets concept and image classification and explores the effectiveness of rough sets theory in image classification by using the change of rough sets theory to improve image classification model. Rough sets theory is used for classification rules extraction.*

Key Words**:** *Rough Set. Change of topologies, Image Classification*

Date of Submission: 11-01-2018 Date of acceptance: 25-01-2018

I. INTRODUCTION

The main problem in the image classification problems and dealing with the huge number of data is the high overlapping between Image classes and uncertainty.

we see the task of extracting the information in the form of patterns, relationships and groupings to be used in applications such as decision support, prediction and taxonomy has become arduous and essential with huge amounts of improper additional texture information may result in a chaotic state, and this leads to uncertainty in the classification process.

Rough set theory is an approach to handle vagueness and uncertainty. We propose methods that apply rough set theory in the context of partitioningof image data . We put this approach to specify the classes in the data by selecting respective regions in 2D slices. Rough set theory compute lower and upper approximations of the classes .The boundary region is the difference between the lower and the upper approximations represents the vagueness of the classification. We present an approach to automatically compute segmentation rules from the rough set classification using expansion of approximation spaces .

This study proposes using Rough Set concepts; depend on the Change of special topology called quasi discrete topology, as a tool for image classification.Weshow that the appropriate classification knowledge can be presented by Rough Set and construct mathematical model to eliminate the huge data, and this information can improve the accuracy of image classification.

1. APPROXIMATION SPACE

Rough sets have been initiated by Pawlak (1982) in order to describe approximation knowledge of subsets of a given universe. In some sense this theory can be considered as a generalization of classical set theory. The description of these sets are derived from an information system. Information system is (U, A) where U is the universe, i.e., a non-empty, finite set of objects, and A is a given of attributes, i.e., a non-empty, finite set of attributes $a \in A$ such that $a: U \rightarrow Va$,where Va denotes the values of attributes that feature a may have. Thus, the information data assigns a value $a(u) \in Va$ to attribute $a \in A$ and object $u \in U$. Information systems are often given in form of a decision rules that lists these values. In the context of image partition, the objects represent pixels and the attributes given the data about pixels such as RGB color. While the decision table contains all pixels, we will in our interactive approach use selected and manually classified pixels as a training set to determine a classification for all pixels.

As such, we introduce another column for the training set, which adds the classification label. This additional row is referred to as the decision feature, while the others are referred to as the conditional features.

This article is devoted to give a study on approximation space. The idea of rough set depends on the concept of approximation space.

Definition 1.1:

The pair A = (U, R) is an approximation space, where U be a finite and R anequivalence relation on U. Pawlak(1991, 1996)

From Definition1.1.1, The subsets of U are definable in the given approximation space and other are undefinable. Thus, a set X⊂U is definable (exact) if it is a union of equivalence classes, otherwise the set X is undefinable.

Now, rough set (undefinable set) can be defined approximately in any approximation space which depend on two exact sets referred to as lower approximation and upper approximation.

Definition 1.2

Given an approximation space A = (U, R) and X⊂U, then

(i) Lower approximation of X is the greatest union of equivalence classes contained in X and denoted by $L(X) = \cup\{[x] \in U/R, [x] \subset X\}$

(ii) Upper approximation of X is the smallest union of equivalence classes containing X and denoted by $u(X) = \cup\{ [x] \in U/R , [x] \cap X \neq \emptyset \}$,

Where U/R is the set of equivalence classes on U. Pawlak (1991, 1996)

Definition 1.3

Given an approximation space A = (U, R) and X⊂U, then

(1) $L_\mu(X)$is the elements in L(X) which is called the internal measureof X.

(2) $uL_\mu(X)$is the elements in u(X) which is called the external measure of X.

(3) $\eta(X)$is the accuracy of X in A and is given by$\eta(X) \frac{L_\mu(X)}{uL_\mu(X)}, uL_\mu(X) \neq 0$.Pawlak (1991, 1996)

DIFFERENT FORMS OF ROUGH SETS

In the following we discuss some different ways to determine rough sets.

(1) Let A =(U,R) given approximation space. Then Set X⊂ U is rough set if L(X) ⊆ X⊆u(X).

From definition 1.3 the rough set satisfies$0 < \eta(X) < 1$.

(2) Let A =(U,R) given approximation space and X⊂U. Then X is

(i) roughly definable if L(X) ≠ ∅ and u(X) ≠U,

(ii) internally roughly undefinable if L(X)= ∅ and u(X) ≠U,

(iii) externally roughly undefinable if L(X) ≠ ∅ and u(X) =U,

(v) totally undefinable if L(X)= ∅ and u(X)=U.Pawlak (1991, 1996)

Remark 1.1(On different Forms of rough sets)

The rough sets is defined in two ways. The first is the accuracy measure and the second is the classification of rough sets. The accuracy coefficient expresses the size of boundary of the set. But it says nothing about the structure of the boundary. While the classification of rough sets gives no information about the size of the boundary region, but it provides us with some information about the boundary region. By the accuracy of a set, we are still unable to tell exactly its topological structure and conversely the knowledge of the topological structure of the set does not gives information's about its accuracy.

Therefore, in practical application of rough sets we combine both kinds of information about the boundary region (accuracy measure) and the information about classification of topology under consideration.

Theorem 1.1

Given an approximation space A = (U, R) and X⊂U, thenthe following statements hold:

(i) X is exact set if and only if the lower approximation and upper approximation coincide.

(ii) X is rough set if the lower and upper approximations are different. Pawlak (1991, 1996)

Definition 1.4

Given an approximation space A = (U, R) and X⊂U, then

The Boundary set of X (denoted by b(X)) is defined by b(X) = u(X) – L(X).

The Internal edge of X is the set of all elements which belong to each of the boundary and the set X and is defined by Int.edge (X) = X – L(X).

The external edge of X is the set of all elements which belong to the boundary set but not in X and is defined byExt.edge (X) = u(X) – X.

The negative region of X is defined by neg.(X)= U-u(X).Pawlak (1991, 1996)

The following figure indicate the above concepts

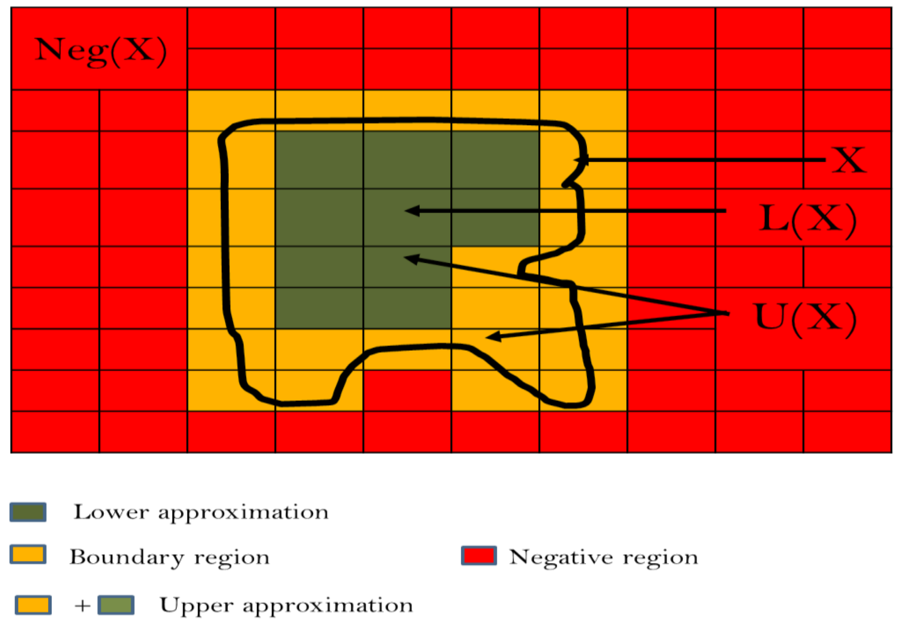

Figure 1

In the above figure, the squares are the equivalence classes. Theset X is the free-hand drawing, the dark gray squares is a lower approximation of X and the dark and light gray squaresisupper approximation of X.the boundary region is the gray squares. The negative region is the white squares.

II. Change of Approximation Space

The aim of this article is to introduce and study the concept of Change of approximation space. Also, some of its properties are investigated.

Definition 2.1
Let A1= (U,R1) and A2= (U,R2) be two approximation spaces. If R1 ≤ R2, then the approximation spaces A1 is finer than A2.

Remark 2.1
If A 1 is finer than A2, then R1 is specialization of R2 (or, R2 is generalization of R1).
Also, if U/R1 ≤ U/R2, then every set of U/R1 is included in a set of U/R2.

Theorem 2.1
Let A1= (U, R1), A2= (U, R2) be two Pawlak approximation spaces and β_1,β_2 be two bases for A 1 and A 2 respectively. If, each B $\in \beta_1$ is the union of members ofβ_2, then A2 is aChange of an approximation space of A1.
Proof: Let G be a composed set of A1.Then G is the union of members of β_1 and hence.

$$G = \bigcup_{i=1}^{n} \{x_i : [x_i] \in \beta_1\}$$

Since, [x]i $\in \beta_1$,then [x]i is the union of members of β_2 , for each I =1,2,3,……,n, hence .

$$G = \bigcup_{i=1}^{n} \{x_{io} : [x_{io}] \in \beta_2\}$$

Therefore, G is composed set in A2, then A1 ⊂ A2.

Theorem 2.2
Let A1= (U, R1), A2 = (U, R2) be two approximation spaces and define the topologies τ_{R1} and τ_{R2} respectively. If R1⊂R2, then $\tau_{R2} \subset \tau_{R1}$ and we say that τ_{R1} is aChange ofτ_{R2}.
Proof: By using Theorem 2.1 and Remark 2.1, the proof is obvious.
Remark 2.2

The bases U/R1 and U/R2 form two partitions of A1 and A2 respectively.

Remark 2. 3

If K = (U,ER) is an expansion of given approximation space A, then the following hold
(i) if X is undefinable (rough) in a given approximation space, then X is also definable (rough) in the expansion of Pawlak approximationspace.
(ii) if X is externally undefinable (internally undefinable and totally undefinable) in a given approximation space, then X is not externally undefinable (is not internally undefinable and not totally undefinable) in expansion of a given approximationspace.
(iii) the expansion lower approximation of X is the greatest union concepts contained in X and denoted by EL(X).
(iv) the expansion upper of X is the least union concepts containing X and denoted by Eu(X).
(v) the accuracy of X in an expansion of an approximation space is given by

$$\eta_E(X)\frac{EL_\mu(X)}{EuL_\mu(X)}, EuL_\mu(X) \neq 0 .$$

Definition 2.2

Let A = (U, R) be a given approximation space and U/R be a partition of U. Then every approximation space has a partition including in U /R which is the Change of A.

Proposition 2.1

If K = (U,ER) is a change of a givenapproximation space of A, then for any X⊂ U the following hold
(i)L(X) ⊂ EL(X)
(ii)Eu(X) ⊂ u(X)
(iii)$\eta \quad (X) \leq \eta_E(X)$

Proposition 2.2

If K = (U, ER) is achange of approximation space of A= (U, R), then the important propertiesof approximation space is like the properties of achange of approximation space.

Example 2.1

Let A = (U, R) be a given approximation space on the set U= {a, b, c, d, e, f, g, h, i, j,k} and U/R={{a, b},{c, g, j}, {d, f},{e, i},{h, k}}. Then the partition of an expansion of approximation space is β_1={{a}, {b}, {c},{g, j}, {d, f},{e, i},{h},{k}}
(i) if X= {a, b, d, f} is definable in U/R, then X is definable in β_1.
(ii) if X= {a, c, d, e, i, k} is externally undefinable in U/R, then X is not externally undefinable in β_1.
Note that, $\eta \quad (X) = 018, \eta_E(X) = 0.71.$
(iii)if X= {c, d, e, i} is rough in U/R, then X is rough inβ_1 Note that,$\eta \quad (X)0.28$, $\eta_E(X) = 060$
(iv) if X= {c, d, e} is internally undefinable in U/R, then X is not internally undefinable in β_1.
Note that,$\eta \quad (X) = 0$,$\eta_E(X) = 0.20.$
(v) if X={a, c, d, e, h} is totally undefinable in U/R, then x is not totally undefinable.
Note that, $\eta \quad (X) = 0,\ \eta_E(X) = 0.50.$

III. Change of Biapproximation Space

The aim of this section is to introduce concept of Biapproximation space based on Pawlak approximation space and give examples to illustrate the behavior of new notion. Before sate the new definition we want mention the following, every element has different view from everyone (mathematician person, chemistries person and ordinary person) and the equivalent classes change from one to other and change when the issues become new. But the problem is when we use the same element (generally, objects) which founded in several situations. So, we try to give new definition help us to use two knowledge bases which have the same objects (or, objects founded in many knowledge bases).

Definition 3.1

If the set U be a finite and R1, R2 equivalence relations on U. Then π = (U, R1, R2) is called Biapproximation space.

Definition 3.2

Let π = (U, R1, R2) be Biapproximation space and define the two Pawlak approximation spaces (or, two knowledge bases) ((U, R1) and (U, R2)) we say that (U, R1) and (U, R2) are equivalent if U /R1 = U /R2 and so biapproximation spaces become (U, R1) or (U, R2), and U /R1 is finer than U /R2, if U /R1 ≤ U /R2(in other hand, (U, R1) is specialization of (U, R2) or (U, R2) is generalization of (U, R1)).

Definition 3.3

Let π = (U, R1, R2) is called Biapproximation space, then the generalization of lower, upper approximation and accuracy measure of A ⊂ $\boldsymbol{U}$with respect to Biapproximation space given by:

$$L_g(A) = L_{R1}(A) \cup L_{R2}(A)$$
$$u_g(A) = u_{R1}(A) \cap u_{R2}(A)$$
$$\mu_g(A) = \frac{|L_g(A)|}{|u_g(A)|}$$

Remark 3.1

1. If U/R = U/R1 ∩ U/R2, then U/R ≤ U/R1 and U/R ≤ U/R2.This means, U/R is specialization of both U /R1 and U /R2).
2. The generalization consists in combining some equivalence classes, where specialization consists in splitting equivalence classes into smaller classes.
3. The Pawlak approximation spaces (U, R1) and (U, R2) enable to indicate the same facts about the universe.
4. If equivalence classes are single elements sets, then the relation is equality relation and give accurate space.

Proposition 3.1

Let π = (U,R1, R2) and R1, R2 are equivalence relations on U. Then the space (U, R) which given by

U/R = {u∩v: u ∈ U/R1;v ∈ U/R2}

forms an Change of both (U, R1) and (U, R2).

Lemma 3.1

Let π = (U, R1, R2) be a bi-approximation space and (U, R) be the Change of (U, R1) and (U, R2). Then the following hold:

1. L π (X) ⊂ L(X)
2. U(X) ⊂Uπ(X

Example 3.1

Suppose we are given the following set of toys blocks, U={x1,x2, x3, x4, x5, x6, x7,x8} and assume that these toys different colors (red, blue, yellow) , shapes (square, round, triangular) and size (small, large) from the following table

Toys	R1=color	R2=shapes	R3=size
x1	red	round	small
x2	blue	square	large
x3	red	triangular	small
x4	blue	triangular	small
x5	yellow	round	small
x6	yellow	square	small
x7	red	triangular	large
x8	yellow	triangular	large

Table 1

Now we calculate the accuracy measure of any set A with respect to all later definition to determine the difference between them.

Let A ={x1,x3, x4,x5}

	L(A)	u(A)	μ(A)
R1	φ	U	0
R2	{x1, x5}	U	.25

R3	φ	{ x1, x3, x4,x5, x6}	0
R1, R2	{x1, x4, x5}	{x1, x3, x4, x5, x7}	.60
R1, R3	{ x1, x3, x4}	{ x1, x3, x4, x5, x6}	.60
R2, R3	A	A	1
R1, R2, R3	A	A	1

Table2

the best generalization of given data is:
U/{R2, R3 }={{x1, x5}, {x2},{x3, x4}, {x6},{x7, x8}}
U/{R1, R2 , R3}={{x1}, {x2},{x3}, {x4},{x5},{x6},{x7},{x8}}.

$L_g(A) = L_{R2}(A) \cup L_{R3}(A) = \{x1, x5\}$

$u_g(A) = u_{R2}(A) \cap u_{R3}(A) = \{x1, x3, x4, x5, x6\}$

$$\mu_g(A) = \frac{|L_g(A)|}{|u_g(A)|} = .40$$

IV. ROUGH SETS AND IMAGE CLASSIFICATION

We applied changes of topologies to the segmentation of image. We have chosen the 2D slice for assembling the training set. We selected few pixels to interactively choose different classes. Given the few samples that have been provided, the results are reasonable. Next, we took the rules we obtained from the first point and assigned them to other points. Segmentation of imaging data is an important and crucial step in the image classification. The rough set analysis requires no external factors and uses the information presented in the given data. In the case of redundant information, only the minimum amount of data needed for the classification model is used. We use the concept of rough sets to express uncertainty in segmentation tasks. The task is to estimate sets of the entire universe that are conform to the attribute values of the training set. These sets can often be described by a few rules to a subset of the attributes.

Rough image is a collection of pixels and the equivalence relation induced equivalence classes and quasi discrete topology.We converted the image into an approximation space by obtaining the equivalence classes through the information about each pixel. We choose eight points of snail see picture of three different parts as following table:

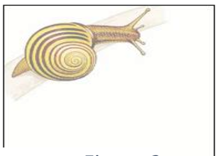

Figure 2

	R	G	B	C
p1	(183.5,Inf)	(-Inf, 249.5)	(-Inf, 132.0)	3
p2	(183.5,Inf)	(-Inf, 249.5)	(132.0,250.5)	2
p3	(183.5,Inf)	(-Inf, 249.5)	(250.5,Inf)	2
p4	(183.5,Inf)	(249.5,Inf)	(250.5,Inf)	1
p5	(183.5,Inf)	(249.5,Inf)	(132.0,250.5)	2
p6	(183.5,Inf)	(249.5,Inf)	(-Inf, 132.0)	3
p7	(-Inf, 183.5)	(-Inf, 249.5)	(132.0,250.5)	3
p8	(-Inf, 183.5)	(-Inf, 249.5)	(-Inf, 132.0)	3

Table 3

For every attribute we construct basis for quasi discrete topology as follows:
U/R = {{p1, p2, p3, p4, p5, p6}, {p7, p8}}
U/G = {{p1, p2, p3, p7, p8}, {p4, p5, p6}}
U/B = {{p1, p6, p8}, {p2, p5, p7}, {p3, p4}}
U/C = {{P1, P6, P7, P8}, {P2, P3, P5}, {P4}}

The base of given data given by: β= {{P1, P6}, {P2, P7}, {P3, P4}, {P5}, {P8}}

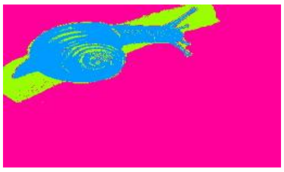

Figure 3

```
if(g<250) {if(b<133) {return C3;} else {return C2;}
          }
else {
                    if(b<133) {return C3;} else {if (b<251) {return C2;}
                       else return C1;
```

by using change of topology concepts, we have a base of new approximation space as follows:
The base of new approximation space given by: β_E= {{P1, P6}, {P2}, {P7}, {P3, P4}, {P5}, {P8}}

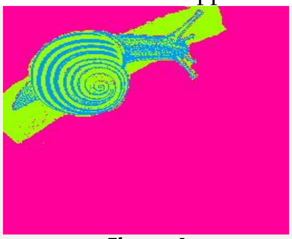

Figure 4

```
if(b<251){
          if(r<184){return C3; } else{ return C2; }
 } else{
          if(g<250){ if(b<133){ return C3; } else { return C2; }
          }     else{
                    if(b<133){ return C3; } else {if (b>251){return C1;}
                       else return C2;
               }
```

V. CONCLUSION

We applied changes of topologies which depend on rough set theory to the segmentation of image. We selected few pixels to interactively choose different classes to construct the rules depend on equivalence class. So, in this paper, we propose a segmentation method that allows us to improve the uncertainty in the data. The segmentation method is based on a classification using change of topologies. One advantage of the rough set theory is the generation of readable if then decision rules as the output of the classification process when applied to a training set. Such rules allow us to reveal new patterns in the data material and to use them as classifiers for unseen data sets.

ACKNOWLEDGEMENTS

I would like to thank Deanship of Scientific Research at Majmaah University for supporting this work.

REFERENCE

[1]. Fibak, Jan, et al. "Rough sets based decision algorithm for treatment of duodenal ulcer by HSV." *Biological Sciences* 34 (1986): 227-249.Hunt, E. B.: Concept formation. John Wiley and Stom, New York, (1974).
[2]. Pawlak, Zdzislaw, and W. Marek. "Rough sets and information systems." *ICS. PAS. Reports* 441 (1981): 481-485.
[3]. Martize, P.: Pawlak and topological rough sets in terms of multifunction, *Glassnik Mathemticki*, vol. 31 (51), 159-178, (1996).
[4]. Nowak, R., R. Slowinski, and J. Stefanowski. "Rough Sets based diagnostic classifier of reducers." *Maintenance Management International* (1990).
[5]. Orłowska, Ewa S. *Logic of vague concepts: applications of rough sets*. Polish Academy of Sciences [PAS]. Institute of Computer Science, 1982.
[6]. Pawlak, Z.: Sets rough relation and rough function,*Fundamental informatica*, 27,103-108, (1996).
[7]. Pawlak, Z.: Rough sets. *ICS PAS Reports* (431), (1981).
[8]. Pawlak, Z.: Rough sets, Algebric and topological approach. *ICS PAS Reports* (482), (1982).
[9]. Pawlak, Zdzisław. "Rough sets." *International Journal of Parallel Programming* 11.5 (1982): 341-356.
[10]. Reinhard, A., B. Stawski, and T. Weber. "Application of rough sets to study the water outflow from the river basin." *Bulletin of the Polish Academy of Sciences, Technical Sciences* 37.1 (1989): 97-104.
[11]. Wiweger, Antoni. "On topological rough sets." *Bulletin of the Polish Academy of Sciences. Mathematics* 37.1-6 (1989): 89-93.
[12]. Yao, Y. Y. "Two views of the theory of rough sets in finite universes." *International journal of approximate reasoning* 15.4 (1996): 291-317.
[13]. Yao, Y. Y. "On generalizing Pawlak approximation operators." *Rough Sets and Current Trends in Computing*. Vol. 1424. 1998.