Decision Support System for Histopathological Diagnosis of HER2 Breast Cancer using Pawlak’s Information System and Mamadani Type Fuzzy Control

Martin Tabakov¹, Krzysztof Rodak¹, Marzenna Podhorska-Okolow², Bartosz Pula², Jedrzej Grzegzolka²
¹Institute of Informatics, Wroclaw University of Technology
Wroclaw, Poland
²Department of Histology and Embryology
Wroclaw Medical University
Wroclaw, Poland
martin.tabakov@pwr.wroc.pl, 164825@student.pwr.wroc.pl, marzenna.podhorska-okolow@am.wroc.pl, {jedrzej.grzegzolka, bartosz.pula}@gmail.com

ABSTRACT: In this article, the specification of a histopathology decision making support system, based on Pawlak’s information system concept and Mamdani type fuzzy control is presented. The proposed system supports the recognition process of HER-2/neu histopathology preparations through microscopy image information analysis. We used Pawlak’s information system to identify the decisive set of features and the optimal set of decision rules under the considered histopathology problem. Then, the so generated decision rules were transformed into fuzzy rules and exploited in Mamdani reasoning. The proposed approach was tested over real clinical data of HER-2/neu breast cancer histopathology images.

Keywords: Rough Sets, Fuzzy Control, Mamdani Reasoning, Breast Cancer, HER-2/Neu Biomarker, Histopathology Image Analysis

Received: 1 November 2012, Revised 10 December 2012, Accepted 15 December 2012

©2013 DLINE. All rights reserved

1. Introduction

In this paper a clinical decision support system, designed to assist physicians and other health professionals with decision making tasks of establishing a diagnosis on the basis of patient histopathology images, is presented.

The main purpose of modern clinical decision support systems is to assist clinicians at certain stage of health care [1]. This means that a clinician would interact with the system to help determine diagnosis, analysis, etc. of patient data. The modern methodology of using clinical decision support system to assist, forces the clinician to interact with the system utilizing both the clinician’s knowledge and the system output to make a better analysis of the patients data than either human or the system could make on their own. In our work, we propose to use rough set methodology, which allows the generation of knowledge base with the assistance of domain experts. What more, we propose an interpretation of the generated knowledge base as a set of fuzzy rules, which enable easy way of applying and analyzing uncertain, vague information. The proposed clinical decision support system was tested on the example of the HER2 breast cancer recognition problem.

Breast cancer is considered as the most common cancer in women and is the first cause of morbidity due to cancer in women age 20 to 59 years. In approximately 20% of the diagnosed breast cancers, an overexpression of Human Epidermal Growth Factor Receptor 2 (HER2/neu) is noted. This receptor is currently recognized as a valuable prognostic and predictive factor related to
the anti-HER2 therapy. Breast cancers characterized by the overexpression of this protein are of higher malignancy grade, proliferative potential and metastatic spread, which in times preceding the utilization of trastuzumab (Herceptin, Genentech, CA; an anti-HER2 agent) treatment led to earlier patients death [10, 12]. In the last years a routine diagnostic procedure based on immunohistochemical (IHC) and fluorescent in situ hybridization techniques (FISH) was established to identify potential responders to trastuzumab therapy, which in numerous studies was shown to reduce the risk of recurrence and mortality rate in early and advanced stage breast cancer [14]. In routine clinical practice, a two-step diagnostic procedure is used for evaluation of HER2 expression status. At first, in IHC HER2 preparations of paraffin-embedded breast cancer specimens a visual semi-quantitative examination of membranous cell staining in tumor cells under a light microscope is performed. For this purpose a categorical classification system based on staining intensity and its pattern is utilized and briefly encoded as follows: 0 (no staining), 1+ (incomplete, weak membrane staining regardless of the proportion of tumor cells stained), 2+ (nonuniform complete membrane staining or staining with obvious circumferential distribution in at least 10% of the tumor cells, or intense, complete membrane staining ≤ 30% of the invasive tumor cells), 3+ (intense membrane staining in >30% of the invasive tumor cells) [14]. For trastuzumab treatment cases scored 3+ in the IHC preparations are classified for the therapy, whereas cases scored 2+ are subjected to further testing with costly FISH examination to finally determine HER2 expression status [3, 14].

So, it is extremely needed to introduce less expensive diagnostic process for correct recognition of the corresponding HER-2/neu classes.

The rough set theory has proven itself highly effective in various medical diagnostic systems enabling reduction of redundant attributes [4, 13, 11]. What’s more, it has also been successfully used in HER-2/neu breast cancer recognition problem along with different classification tools. In [9] support system based on rough sets and a probabilistic neural networks is proposed. Authors report effective reduction of 67% redundant attributes from a biomedical dataset without sacrificing classification accuracy results. Furthermore, fully automatic method for breast cancer diagnosis, which utilize rough sets and multi-category proximal support vector machine (MPSVM) is presented in [5]. Feature extraction is based on wavelet moments, which results in generating almost 800 features per image, hence dimensionality reduction is necessary in order to perform efficient classification. Scores shows around 90% accuracy against visual expert examination, which also proves rough sets robustness.

In this work, we propose histopathology decision support system that uses rough sets to generate the optimal set of decision rules and next in order to refer the fuzzy aspect of the considered histopathology problem, we transform the generated rules into fuzzy rules. The last enable the use of fuzzy reasoning methods (we have focused on the classical Mamdani reasoning) as a decision support methods.

Thus, using rough sets and what more combining this approach with fuzzy control, we proposed effective methodology for histopathology decision support.

The article is organized as follows: in section 2 some theoretical background, considering the methods used in the proposed research, are described; in section 3, the suggested decision support process is introduced and finally, in section 4 and 5 some experiments and results discussion are presented.

2. Theoretical background

In this section, the preliminaries of fuzzy sets [15], fuzzy control of Mamdani type [6] and rough sets [8] are described.

2.1. Fuzzy sets

Let $\mathcal{X} = \{x_1, x_2, ..., x_n\} \subseteq \mathbb{R}$ be some finite set of elements (domain), then we shall call ‘$A$’ the fuzzy subset of $\mathcal{X}$, if and only if:

$A = \{ (x, \mu_A(x)) \mid x \in \mathcal{X} \}$, where $\mu_A$ is a function that maps $\mathcal{X}$ onto the real unit interval $[0, 1]$, i.e. $\mu_A : \mathcal{X} \rightarrow [0, 1]$. The function $\mu_A$ is also known as the membership function of the fuzzy set $A$, as its values represents the grade of membership of the elements of $\mathcal{X}$ to the fuzzy set $A$. Here the idea is that we can use membership functions, as characteristic functions (any crisp set can be defined by its characteristic function) for fuzzy, imprecisely described sets. Let $A$ and $B$ be two fuzzy subsets of $\mathcal{X}$, then the basic set operations: union and intersection of $A$ and $B$, are defined as follows:

$\mu_{A \cup B}(x) = \max \{ \mu_A(x), \mu_B(x) \}$, $\mu_{A \cap B}(x) = \min \{ \mu_A(x), \mu_B(x) \}$.
2.2. Fuzzy Control with Mamdani reasoning

We propose to use fuzzy control concept, as a decision mechanism in our decision support system. For any fuzzy controller, all input information are fuzzified and then processed with respect to assumed knowledge base, inference method and defuzzification process.

A fuzzy controller is composed of the following four elements:

- A rule-base (a set of IF-THEN rules), which contains a fuzzy logic quantification of the expert’s linguistic description of how to achieve good control.

- An inference mechanism, which emulates the expert’s decision making in interpreting and applying knowledge.

- A fuzzification interface, which converts controller inputs into information that the inference mechanism can easily use to activate and apply rules.

- A defuzzification interface, which converts the conclusions of the inference mechanism into actual inputs for the process.

More formal approach, assuming Mamdani reasoning, is given below.

Let \( \{ X_1, X_2, \ldots, X_n \} \) be a family of finite sets, which defines the corresponding states of primary inputs for a fuzzy control system. Let \( V(X_1^i) = \{ V^1_{X_1^i}, V^2_{X_1^i}, \ldots, V^n_{X_1^i} \} \) be a set of all linguistic variables defined by experts for \( X_i \), i.e. fuzzy subsets of \( X_i \), where \( i = 1, 2, \ldots, n \). Then a fuzzy IF-THEN rule antecedent takes the form: \( (x_1 \text{ is } V^1_{X_1^i}) \circ (x_2 \text{ is } V^2_{X_1^i}) \circ \ldots \circ (x_n \text{ is } V^n_{X_1^i}) \), where \( (x_i \text{ is } V^i_{X_1^i}) \) means the degree of membership of \( x_i \) to \( V^i_{X_1^i} \) \( (x_i \in V^i_{X_1^i}) \), \( i = 1, 2, \ldots, n \). This suggests that a fuzzy control rule antecedent is interpreted as a weight coefficient that represents the strength of firing the rule.

Let consider a simplified example fuzzy rule:

\[
\text{Rule} = \text{IF } (x_1 \text{ is } V^1_{X_1^i}) \circ \ldots \circ (x_n \text{ is } V^n_{X_1^i}) \text{ THEN } (\text{conclusion } = Y^p_{k_1} \circ \ldots \circ Y^p_{k_m})
\]

The antecedent value of the rule is interpreted as a weight coefficient that represents the strength of firing the rule.

Let consider a simplified example fuzzy rule: \( r: \text{IF } (x \text{ is } A) \text{ THEN } (y \text{ is } B) \), where \( A \) and \( B \) are some fuzzy sets. In essence, the expression describes a relation between two variables \( x \) and \( y \), such that \( (x, \mu_A(x)) \in A \) and \( (y, \mu_B(y)) \in B \). This suggests that a fuzzy rule can be defined as a binary fuzzy relation \( \rho \) on the product space \( A \times B \), i.e. \( \rho: A \times B \rightarrow [0, 1] \). Based on the interpretations of the Cartesian product and various t-norm and t-conorm operators, a number of qualified methods can be formulated to calculate \( \rho \), as it can be viewed as a fuzzy set with a two-dimensional membership function: \( \mu_{\rho}(x, y) = df \int f(\mu_A(x), \mu_B(y)) \), where the function \( f \), called the fuzzy implication function, performs the task of transforming the membership degrees of \( x \) in \( A \) and \( y \) in \( B \) into those of \( (x, y) \) in \( A \times B \). In this paper, we used the Larsen fuzzy implication function, defined as follows: \( f(\mu_A(x), \mu_B(y)) = \mu_A(x) \cdot \mu_B(y) \), where \( \cdot \) is the algebraic product operator.

Next, in our experiments, we used the Mamdani inference mechanism and the center of gravity (in short COG) defuzzification method. Thus, using the above designations, we can define the final fuzzy control system output value (the decision value) as follows:

\[
\mu = \frac{\int_{Y^p_{k_1} \cup Y^p_{k_2} \cup \ldots \cup Y^p_{k_m}}^y \mu_{\rho} \cdot \mu_{\psi} \cdot \mu_{\tau} \cdot \mu_{\chi}}{\int_{Y^p_{k_1} \cup Y^p_{k_2} \cup \ldots \cup Y^p_{k_m}} \mu_{\psi} \cdot \mu_{\tau} \cdot \mu_{\chi}} (1)
\]

where:
y \in Y; Y'_1, Y'_2, ..., Y'_{m_y} - denotes the possible fuzzy rules conclusions modified by the corresponding rule antecedent values, i.e.: let \( w_r \) be the antecedent value of the \( r \)-th fuzzy rule and let denote the conclusion of this rule as \( Y'_r \). Then, \( Y'_r \) is the corresponding fuzzy set generated from \( Y'_r \) with respect to the \( w_r \), such that (assuming the Larsen fuzzy implication function) \( \mu_{Y'_r}(y) = \mu_{W_r}(y) \) (assuming the Larsen fuzzy implication function).

2.3. Rough sets
A rough set is interpreted as a formal approximation of a crisp set, by a pair of sets, which give the so called lower and upper approximation of the original set. Let consider the classical Pawlak's information system: \( IS = (U, A, V, f) \), where: \( U \) is some universe, \( A \) is a set of attributes, \( V \) is the attributes domain set \( V = \bigcup_{a \in A} V_a \), \( V_a \) - is the domain of the \( a \)-th attribute \(( a \in A )\) and \( f \) is the information function \( f: U \times A \rightarrow V; \forall x \in U, a \in A f(u, a) \in V_a \). Regarding to the following equivalence relation: \( IND(B) = \{ (x, y) \in U \times U; \forall a \in B f(x, a) = f(y, a) \} \), where \( B \subseteq A \), the lower and the upper approximation of a subset of \( U \) can be introduced as follows:

\[
\begin{align*}
B \downarrow X &= \{ x \in U; [x]_{IND} \subseteq X \}, \\
B \uparrow X &= \{ x \in U; [x]_{IND} \cap X \neq \emptyset \}, \text{where } X \subseteq U \quad (2) \\
\text{and } B \subseteq A. \\
\end{align*}
\]

The above information system can be interpreted and realized as a classical decision table (assuming a decision attribute). Using the mathematical apparatus defined for rough sets, there is possible to identify the reduct and the optimal set of decision rules, derived from a given decision table. The reduct gives the minimal set of attributes that fully characterize the knowledge represented in the equivalence class structure. Next, over the derived reduct, the minimal set of decision rules that cover the corresponding decision problem, can be generated. Also, if there is data with conflict objects - i.e. two objects are conflicting when they are characterized by the same values of all attributes, but they belong to different decision classes, the lower and the upper approximation precision can be used to eliminate decision table inconsistency.

In our work, we have used the rough sets to extract optimal set of decision rules, in the following sequence:

a) Design a decision table, regarding to the considered problem (identify the set of objects, attributes and decision classes),

b) Eliminate object conflicts, using the lower approximation precision: let \( X \subseteq U \) \((X \neq \emptyset)\) and \( B \subseteq A \) \((B \neq \emptyset)\), then the lower approximation precision of the set \( X \) regarding to the subset \( B \) is defined as follows:

\[
\gamma_B(X) = \frac{|B \downarrow X|}{|U|} \quad (4)
\]

c) Generate the attributes reduct: generate discernibility matrix \( M(IS) \) - discernibility matrix for information system \( IS: M(IS) = \begin{bmatrix} m_{ij} & \ldots & m_{ij} \end{bmatrix} = \{ a \in A; f(x, a) \neq f(x, a) \} \), where \( n = |U| \) and next apply the Johnson heuristic algorithm for rough set reduction, applied to find single reduct (subset of attributes) \([7]\),

d) Generate the minimal set of decision rules that correctly cover the decision problem (see algorithm 1, below).

Algorithm 1 (rule extraction):

To generate the optimal set of rules over a decision table, the following steps should be taken:

Step 1: Generate the ‘\( M_k \)’ matrixes (matrixes derived over M(IS), which are used to define the so called implications of the considered objects) from the discernibility matrix,

Step 2: Define the object implicants,

Step 3: Define the target set of rules.

Step 1:
\((k = 1, ..., n; n = |U|)\) Let \( e_y \) are the elements of \( M(IS) \), \( \hat{c}_y \) are the elements of \( M_k \) (with respect to \( k \)) and \( a^* \) is the decision attribute.
Then:

For each $k = 1, \ldots, n$:

1) If $i \neq k$ then $\hat{c}_{ij} = \emptyset,$

2) If $(c_{jk} = \emptyset)$ and $(d_{a}^{B} (x_{j}) \neq \{a \ast (x_{j})\})$ then $\hat{c}_{ij} = d_{a}^{c_{jk} \cap B}$ else $\hat{c}_{ij} = df \emptyset$

where $B \subseteq A$ and $d_{a}^{B} (x_{k}) = df \{v \in V_{a} \mid \exists y \in U (x_{i} \text{ IND } y) \land a \ast (y) = v\}$

**Example 1:**

Let consider the following decision table (extended with the corresponding $d_{a}^{A}$ values; $A = B = \{a, b, c\}$):

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>a*</th>
<th>$d_{a}^{A}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>${0}$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>${1}$</td>
</tr>
<tr>
<td>$x_3$</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>${0}$</td>
</tr>
<tr>
<td>$x_4$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>${2}$</td>
</tr>
<tr>
<td>$x_5$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>${0}$</td>
</tr>
</tbody>
</table>

The corresponding discernibility matrix (it has the symmetric property) takes the form:

<table>
<thead>
<tr>
<th></th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$\emptyset$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_2$</td>
<td>${a, c}$</td>
<td>$\emptyset$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_3$</td>
<td>${a}$</td>
<td>${a, c}$</td>
<td>$\emptyset$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_4$</td>
<td>${a}$</td>
<td>${c}$</td>
<td>${a}$</td>
<td>$\emptyset$</td>
<td></td>
</tr>
<tr>
<td>$x_5$</td>
<td>${b}$</td>
<td>${a, b, c}$</td>
<td>${a, b}$</td>
<td>${a, b}$</td>
<td>$\emptyset$</td>
</tr>
</tbody>
</table>

Next, applying the Algorithm 1 for $k = 1$, we can generate the matrix $M_{1}$:

<table>
<thead>
<tr>
<th></th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$\emptyset$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_2$</td>
<td>${a, c}$</td>
<td>$\emptyset$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_3$</td>
<td>$\emptyset$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

as:

$d_{a}^{A} (X_{2}), d_{a}^{A} (X_{4}) \neq \{0\} (j = 2, 4), \text{ thus } \hat{c}_{ij} = c_{ij} \cap A (j = 2, 4), \text{ so: } \{a, c\} \cap A = \{a, c\}, \{a\} \cap A = \{a\}$

Similarly, we can generate the matrixes: $M_{2}, M_{3}, M_{4}, M_{5}$.

Step 2:

Next, we can determine the set of 'object implicants' from each matrix:

**Implicant 1**: from $M_{1}$: $x_{1} \Rightarrow (a \lor c) \land a$ (we can simplify, assuming Boolean algebra and using the corresponding Boolean algebra reduction rules: $x \land x = x \land x; x \land (x \land y) = x$ and $x \lor (x \land y) = x$; $x1 \Rightarrow a$,

**Implicant 2**: from $M_{2}$: $x_{2} \Rightarrow c$.
Implicant 3: from $M_3 : x_3 \Rightarrow a$

Implicant 4: from $M_4 : x_4 \Rightarrow a \land c$

Implicant 5: from $M_5 : x_5 \Rightarrow a \lor b$

(intuitively, the object implicants can be considered as indication concerning which attributes are strongly related to which objects).

Step 3:

Finally, using the above implicants, we can generate the target set of rules, derived from the considered decision table. Each rule concern one decision value and it is derived as a sum of the object implicants related to that decision, i.e.:

Rule 1 (for decision value 0): $f(x_1, a*) = f(x_3, a*) = f(x_5, a*) = 0): f(x_4, a) \lor f(x_5, a) \lor (f(x_5, a) \lor f(x_5, b)) \Rightarrow \text{(decision: 0),}$

Rule 2: $f(x_2, c) \Rightarrow \text{(decision: 1)},$

Rule 3: $f(x_4, a) \land f(x_4, c) \Rightarrow \text{(decision: 2)}.$

3. Method specification

The clue of the decision support method presented in this paper, is our proposition to combine rough sets and fuzzy control, by interpreting the corresponding features, derived from HER-2/neu histopathology images, as domains over which fuzzy sets are defined. This allows to use the fuzzy control concept (in this case the Mamadani reasoning), to generate decision values. The basics of the proposed theoretical conception is presented on the scheme 1, below.
For better explanation of the above scheme, let consider the next example.

**Example 2:**
Let suppose that we have the following decision table:

<table>
<thead>
<tr>
<th>feature&lt;sub&gt;1&lt;/sub&gt;</th>
<th>feature&lt;sub&gt;2&lt;/sub&gt;</th>
<th>decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>small</td>
<td>medium</td>
<td>D&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>medium</td>
<td>small</td>
<td>D&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>large</td>
<td>small</td>
<td>D&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

where, \( U = \{ \text{pattern}_1, \text{pattern}_2, \text{pattern}_3 \} \) – the term ‘pattern’ can be considered as an input information source (for instance – it can be defined as a sub-region of a histopathology image – a tumour cell for example); \( \{ \text{feature}_1, \text{feature}_2 \} \) is the set of extracted image features; \( \{ \text{small}, \text{medium}, \text{large} \} \) are some fuzzy subsets defined over the domains of the considered image features and \( D_1, D_2 \) are some decisions – fuzzy sets, modelling the fuzzy rules conclusions, regarding to the Mamdani model.

Then, we propose the following interpretation of information functions: \( f(\text{pattern}_1, \text{feature}_1) = \text{‘small’} \Rightarrow \text{df ‘feature}_1\text{ of pattern}_1\text{ is ‘small’} \) (the degree of membership of \( \text{feature}_1 \), derived from \( \text{pattern}_1 \), to fuzzy set ‘small’). The used fuzzy sets are defined over the domains of the assumed features, with respect to all considered patterns. What more, we proposed to determine the values of the used information function, regarding to the corresponding membership values, i.e. \( f(\text{pattern}_1, \text{feature}_1) \) is valued as ‘small’, because it is assumed to give highest membership value:

\[
\mu_{\text{small}}(X_{\text{pattern}_1}) > \max\{\mu_{\text{medium}}(X_{\text{pattern}_1}), \mu_{\text{large}}(X_{\text{pattern}_1})\}, X \in \text{feature}_1
\]

Next, let suppose that we have generated the following set of decision rules over the above decision table:

- \( f(\text{pattern}_1, \text{feature}_1) \lor f(\text{pattern}_2, \text{feature}_1) \Rightarrow \text{decision: } D_1 \),
- \( f(\text{pattern}_3, \text{feature}_1) \land f(\text{pattern}_3, \text{feature}_2) \Rightarrow \text{decision: } D_2 \).

Thus, we propose the following interpretation of the above rules in terms of fuzzy rules:

- IF (\( \text{feature}_1 \) is ‘small’) \( \oplus \) (\( \text{feature}_1 \) is ‘medium’) THEN (decision is \( D_1 \)),
- IF (\( \text{feature}_1 \) is ‘large’) \( \otimes \) (\( \text{feature}_2 \) is ‘small’) THEN (decision is \( D_2 \)).

where \( \otimes, \oplus \) are assumed as the Zadeh’s triangular norms.

Once, if we derive the set of fuzzy rules that correspond to the considered histopathology problem, we are able to apply the Mamdani reasoning model and to generate the corresponding decision support values. Of course, it should be noticed that in our system, the generation of the appropriate decision table is being done with cooperation of medical experts.

For better understanding of the proposed concept, below sample model example is considered.

Let suppose that we want to sort a set of images with respect to a certain colour criteria – a set of violet/blue images of flowers with respect to the degree of ‘violet colour’. It is assumed that the fuzzy term ‘very violet’ corresponds to ‘dark-blue’ colours and the fuzzy term ‘low violet’ corresponds to ‘bright-violet’ colours. The goal is to find the most ‘violet’ images. Under these assumptions, the below decision table can be introduced for a selected learning set.

We use the following fuzzification concept: for every image feature over the considered learning set, a Gaussian distribution was generated, which was used to define the fuzzy set ‘medium’ over each image feature: \( \mu_{\text{medium}}(x) = \frac{-\alpha}{\sigma^2}, \) where \( x_0 \) is the expected value and \( \sigma \) is the standard deviation.
Next, the fuzzy sets ‘low’ and ‘high’ were defined as follows:

\[
\mu_{low}(x) = \begin{cases} 
\frac{-1}{1 - e^{2\sigma^2}} & : x < x_0 \\
0 & : x \geq x_0
\end{cases} \\
\mu_{high}(x) = \begin{cases} 
\frac{-1}{1 - e^{2\sigma^2}} & : x \geq x_0 \\
0 & : x < x_0
\end{cases}
\]

also, for example, \(f(image_1, feature_1) = 'medium',\) because \(\mu_{medium}(x_{image_1}) > \max \{\mu_{low}(x_{image_1}), \mu_{high}(x_{image_1})\}, x \in \text{feature}_1.\)

The decision attributes fuzzification can be completed in a classical manner – introducing fuzzy sets that corresponds to the ‘degree of decision’, how ‘violet’ a certain image is. Next, applying the proposed concept, it is possible to generate a numerical values, according to the Mamdani fuzzy model, which corresponds to the input recognition criteria (see table, below).

<table>
<thead>
<tr>
<th>Image</th>
<th>Feature_1 (mean S, HSV colour model)</th>
<th>Feature_2 (mean R, RGB colour model)</th>
<th>Feature_3 (mean G, RGB colour model)</th>
<th>Feature_4 (mean [B–R], RGB colour model)</th>
<th>Decision attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image_1</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Average Violet</td>
</tr>
<tr>
<td>Image_2</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>Low Violet</td>
</tr>
<tr>
<td>Image_3</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Very Violet</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>Image_k</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>

As it can be noticed, classification process by thresholding can be introduced.

4. Experiments and results

In purpose to adjust the proposed method to HER2 histopathology decision support, we have executed the following steps:

• **Define common input for the system** – in the presented system, the input is a set of histopathology image fragments (100 × 100 pixels), derived from histopathological preparation. What more, only the fragments with high entropy value have been considered. The last gives the potentiality to avoid any irrelevant image region, such as: background, unimportant cells and so on (see figure 1, below).
Define learning set – a set of histopathology image fragments, including corresponding decision values, chosen by experts. The decision values used, were interpreted as fuzzy sets (modelling the fuzzy rules conclusions, regarding to the Mamdani fuzzy control model), defined by standard triangular functions which reflect the degree of malignancy of cancer cells (HER2 overexpression). Four decision values were used: $P_1, \ldots, P_4$.

Identify the corresponding feature set for any fragment – in our experiments we have used only basic colour values, derived over the HSV and RGB colour models,

Define fuzzification of the considered image features – define fuzzy sets: small, medium and large over any image feature. To complete this stage, we have divided the available HER2 image dataset into three major classes, refereeing to ‘low aggressiveness cancer cells’, ‘medium aggressiveness cancer cells’ and ‘high aggressiveness cancer cells’. Then we have assumed membership functions of the fuzzy sets used: ‘small’, ‘medium’ and ‘large’ as Gaussian probability density function for any image feature with respect to the generated cancer cells aggressiveness classes, i.e. fuzzy set ‘small’ for any image feature is defined as Gaussian probability density function over all images of the ‘low aggressiveness cancer cells’ with respect to corresponding image feature,

Define corresponding decision table – if we have learning set (HER2 image fragments, image features with corresponding fuzzification and decision values), we are able to build decision table for the considered histopathology problem,
• **Apply rough sets** – eliminate object conflicts, generate attributes reduct and finally, optimal set of decision rules,

• **Apply fuzzy control** – transform the generated set of rules into fuzzy rules and apply fuzzy control inference mechanism,

• **Define common output of the system** – interpretable as a decision making support.

Below, some system action results are shown and the system output interpretation is given (having regard to the limitations of the size of the histopathology images, only chosen HER2 image fragments are considered).

<table>
<thead>
<tr>
<th>System Input (HER2 image fragment):</th>
<th>System Input:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image 1" /></td>
<td><img src="image2.png" alt="Image 2" /></td>
</tr>
<tr>
<td>System output (fuzzy control output value – see equation (1) : 0.2)</td>
<td>System output: 0.4382932354636327</td>
</tr>
<tr>
<td><img src="image3.png" alt="Image 3" /></td>
<td><img src="image4.png" alt="Image 4" /></td>
</tr>
<tr>
<td>System output: 0.3999999999999997</td>
<td>System output: 0.7</td>
</tr>
<tr>
<td><img src="image5.png" alt="Image 5" /></td>
<td><img src="image6.png" alt="Image 6" /></td>
</tr>
<tr>
<td>System output: 0.6</td>
<td>System output: 0.6</td>
</tr>
</tbody>
</table>

**Results interpretation:** high output values refer to high cancer HER overexpression and thus appropriate for trastuzumab treatment. The HER2 overexpression is related (in terms of histopathology HER2 image information) to darker cell membrane staining.

So, the HER2 overexpression is related to high system output values, which is very important information for physicians. This can be also used for decision making in quantitative manner, i.e. *IF* most of the fragments values of a considered HER2 image are \( \geq T \) (where \( T \) is a priori given threshold value), *THEN* classify as appropriate for trastuzumab treatment *ELSE* classify as not appropriate for trastuzumab treatment.

Using the above rule, we have tested our system under real histopathology data (above 72 HER2 images with a priori given FISH examination results for the corresponding histopathology preparations) and different threshold values.

In our best result (the most optimal threshold value), we have achieved 82.6% of system accuracy.
What more, as the compactness of the cell membrane is very important for the recognition of the HER2 over-expressed cancer cells, we also tested the proposed method over the cell membranes (detected by standard edge detectors) with respect to chosen shape coefficients. We have used coefficients, which are directly related to the compactness of the considered cell membranes (considered below as objects), such as: Blair-Bliss coefficient: $\frac{S}{\sqrt{2\pi} \sum_i r_i}$ (where $S$ – is the object area, $r$ – is the distance of $i$th object pixel to the object’s gravity center), Feret coefficient: $\frac{L_h}{L_v}$ (where $L_h$ - circuit object in horizontal direction, $L_v$ - circuit object in vertical direction), box counting dimension (fractal dimension analysis procedure), shape circularity: $\frac{2\sqrt{S}}{\pi}$, and other.

Below, some system action results over the shape coefficients of the cell membranes (assuming also skeletonisation) are shown and the system output interpretation is given (having regard to the limitations of the size of the histopathology images, only chosen HER2 image fragments are considered).

<table>
<thead>
<tr>
<th>Image</th>
<th>System output:</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td>0.553373589006844</td>
</tr>
<tr>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td>0.558075184987646</td>
</tr>
<tr>
<td><img src="image3.jpg" alt="Image 3" /></td>
<td>0.38733690002583</td>
</tr>
<tr>
<td><img src="image4.jpg" alt="Image 4" /></td>
<td>0.345491002096553</td>
</tr>
</tbody>
</table>

Results interpretation: high output values refer to high cancer HER overexpression (in terms of shape coefficient, it means high membrane compactness) and thus appropriate for trastuzumab treatment. The HER2 overexpression is related (in terms of histopathology HER2 image information of membrane shape coefficients) to more compact cell membranes.
Therefore, it is possible, to generate one final result, by aggregation of the obtained results over different image characteristics (colour features, shape coefficients) using t- and s- norms [2] as aggregation operators (providing corresponding normalization process to the [0,1] real number interval).

5. Conclusions

The contemporary for breast cancer diagnosis is visual examination of microscopic biopsy images. Nevertheless, manual expert classification is both inefficient and ineffective, hence more robust and automatic approach is needed. In this study, we present histopathology decision making support system, based on rough sets and fuzzy control for HER-2/neu image analysis.

In presented solution, preprocess phase enables to focus on the most crucial parts of biopsy image, using entropy value to difference each image segment. Applying fuzzy sets to features vector helps to reflect fuzzy characteristic of image attributes, providing more appropriate description then crisp value. The rough sets successfully helped to reduce features vector and remove conflicting objects from dataset, providing foundation for more efficient classification. Lastly, fuzzy rules and fuzzy control system along with Mamdani reasoning leads to precise classification. System accuracy is estimated as 82% concordance with conclusive FISH test.

The proposed system was developed in purpose to provide additional input for experts, which may increase their general accuracy.

The presented work may be improved in several ways. Firstly, current feature vectors contains only color-related and shape related attributes. It can be extended by adding any other characteristics of cell membranes or image information (texture analysis for example), which should help to enhance the considered feature vectors. Secondly, more sophisticated methods of extracting important fragments of biopsy images may be used, along with more advanced reasoning mechanism or adding type-2 fuzzy sets to the fuzzification process. Finally, the dataset used in this research should be enlarged to provide more accurate results, which would be essential to test presented approach before using the system in clinical settings.

6. Acknowledgment

This work was partially supported by Polish Ministry of Science and Higher Education grant No N N518290540, 2011-2014.

References


