

# USING RAPIDMINER FOR SEMIAUTOMATIC CLASSIFICATION OF JAWBONE CYSTS FROM RTG SNAPSHOTS

**Lenka Polášková**

Bachelor Degree Programme (2), FEEC BUT

E-mail: xpolas28@stud.feec.vutbr.cz

Supervised by: Jan Mikulka

E-mail: mikulka@feec.vutbr.cz

**Abstract:** In cases when a classification of the objects by a person is not feasible or is not appropriate, there is a possibility of using a computer system which consists of special tools for these purposes. In order to facilitate the work with more than one tool it is crucial to interconnect the processes running in respective tools. The task is to develop a component connecting an image processing and a data mining provided by two different special environments. This paper presents a method for semiautomatic evaluation of cystic disorders in human jawbones taking benefits of mentioned approach. Apart a medicine support, the aim of this project is also to create an instrument offering a draft for solving similar problematic.

**Keywords:** ImageJ, image processing, RapidMiner, data mining, jawbone cysts

## 1. ÚVOD

Pro každého lékaře vychází diagnostika čelistní cysty z jeho znalostí a zkušeností, je tedy subjektivní. Diagnostika cysty z rentgenového snímku je však více subjektivní a sporná, než diagnostika této cysty například histologickou analýzou. Tento projekt by měl lékařům usnadnit použití vyhodnocovacích algoritmů pro diagnostiku dvou typů čelistních cyst. Jedná se o cysty radikulárního a folikulárního typu. Cílem této práce je poloautomatické vyhodnocení typu čelistní cysty, přičemž výsledkem je informace o tom, zda navržený systém vyhodnotil cystu jako cystu prvního typu nebo jako cystu druhého typu.

## 2. ZPRACOVÁNÍ OBRAZU

Řešení projektu vyžaduje znalosti z oblasti zpracování obrazů. Zpracování obrazů zahrnuje kroky uvedené na Obr. 1.



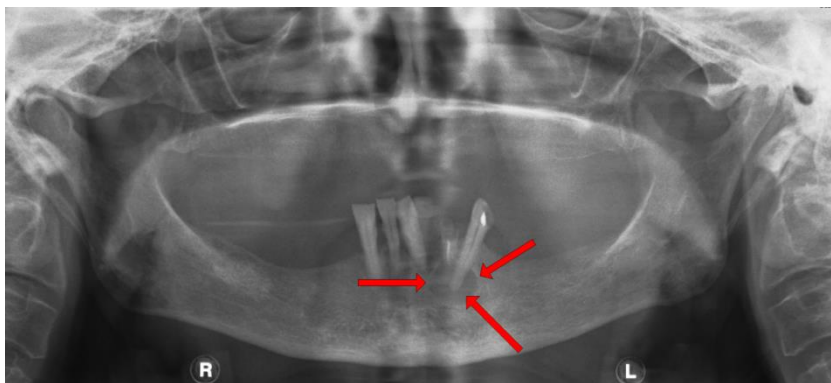
Obr. 1: Řetězec zpracování obrazu

Rentgenové obrazy byly pořízeny a předzpracovány ve Fakultní nemocnici v Brně. Další předzpracování a následná segmentace, popis a klasifikace patří již k oblastem, do nichž je zasaženo v rámci řešení projektu. Předzpracování obecně zahrnuje metody, které je možné použít pro zlepšení kvality obrazu, zvýšení poměru signálu k šumu a zvýraznění některých rysů pro lepší reprodukovatelnost objektů v obraze. Výsledkem je obraz upravený dle účelu dalšího zpracování. Po předzpracování je potřeba v obraze vybrat (ohraničit) oblast zájmu – cystu. Nabízejí se různé způsoby segmentace pro nalezení oblasti zájmu v obraze, avšak pouze dvě z těchto metod poskytují dobré výsledky při ohraničování čelistních cyst. První z nich je provedena manuálním ohraničením, kdy je

oblast zájmu nalezena spojnicí vykreslenou samotným uživatelem, a druhá z nich, metoda live-wire, využívá množinu bodů navržených uživatelem a Dijkstrova algoritmu pro hledání nejbližší cesty mezi dvěma body. Jakmile je cysta ohraničena, následují kroky popisu a klasifikace cysty, které budou popsány v dalších kapitolách podrobněji.

### 3. ČELISTNÍ CYSTY

Cysta je ve tkáni vytvořená dutina s vlastní epiteliární výstelkou. Může mít epiteliární obal oddělující vlastní obsah od okolního parenchymu. Cysta naplňuje svůj objem metabolity a zvětšuje se. Kvůli zvětšování dochází v okolí cysty k utlačování struktur a jejich odbourávání [1]. Na Obr. 1 Obr. 2 je uveden panoramatický rentgenový snímek čelistí s cystou ležící ve frontální oblasti dolní čelisti.



Obr. 2: Červené šipky poukazují na cystu radikulárního typu

#### 3.1. CYSTY RADIKULÁRNÍ

Radikulární cysty patří do kategorie zánětlivých cyst. Vznikají zánětlivým drážděním proliferací zbytků Malassezových epiteliárních buněk po nekróze zubní dřevě v důsledku zubního kazu nebo traumatu. Radikulární cysty nasedají periapikálně na kořen zubu, apex může být patrný v lumen cysty. Nachází se typicky ve frontálním úseku čelisti [1].

#### 3.2. CYSTY FOLIKULÁRNÍ

Cysty folikulární patří do kategorie vývojových. Vyvíjejí se mezi vrstevmi sklovinného epitelu a/nebo mezi sklovinným epitelem a zubní korunkou nadpočetného nebo regulárního zubu. Folikulární cysty na rentgenovém snímku nasedají na cementosklovinnou hranici a je pro ně typické, že korunka zubu nebo její část je umístěna v lumen cysty. U velkých cyst nemusí být tato topografie dobře patrná. Cysta roste pomalu a přitom deformuje čelist. Většinou jsou jednodukomorové, ale mohou být i vícekomorové. V čelisti se nacházejí typicky laterálněji [1].

Ke klasifikaci bylo použito 37 cyst. Cysty byly klasifikovány lékaři tak, že 25 cyst bylo kategorizováno do typu radikulárních cyst a 12 cyst bylo kategorizováno do typu folikulárních cyst.

### 4. IMAGEJ

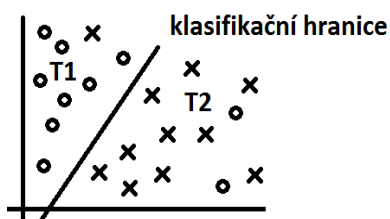
Program ImageJ poskytuje řadu funkcí pro zpracování obrazu (*image processing*). ImageJ je volně stažitelný program, vyvinutý společností National Institute of Health. Je navržen v Javě a jeho architektura je otevřená, takže je možné jej rozšiřovat pluginy a makry. V tomto programu je rentgenový obraz segmentován a dále je prováděna jeho analýza. Analýza sestává z měření charakteristických rysů vybrané oblasti zájmu. Hledány jsou minimální a maximální hodnoty šedé barvy, jejich rozdíl, extensivní hustota šedé barvy a hodnoty x-ových souřadnic středu a těžiště cysty. Tyto hodnoty jsou využity při klasifikaci čelistních cyst. Navržený doplněk pro klasifikaci cyst se stane součástí tohoto programu.

## 5. RAPIDMINER

RapidMiner je světově rozšířený *open-source* program vyvinutý společností Rapid-I určený pro strojové učení (*machine learning*) a dolování dat (*data mining*) a využíváný ve více než 40 státech světa a v mnoha velkých společnostech jako je Ford, Nokia, IBM, Honda, E.ON, atd. Je také navržen v Javě, takže ho taktéž lze spustit jako stažitelnou aplikaci, nebo pouze na *java virtual machine* bez instalace. Díky možnosti spuštění bez instalace pro lékaře odpadá povinnost stahování a instalace programu. Rapidminer využívá pro popis struktury procesu jazyk XML. Tento jazyk je totiž široce využívám právě pro popis struktury objektů a výměnu dat. Rapidminer podporuje práci s databázemi [2] jako je CVS, excel, MS Acces, Oracle a mnoho dalších. V tomto projektu bude využita jedna z důležitých funkcí tohoto programu – schopnosti učení. RapidMiner má inteligentní algoritmy pro učení z dané databáze známých dat pro pozdější predikci dat neznámých. Součástí procesu učení je klasifikace vstupních dat a validace klasifikačních výsledků zahrnující statistickou analýzu úspěšnosti klasifikace. Jako vstupní data zde poslouží parametry obrazu cyst naměřené v programu ImageJ a uložené do databáze.

### 5.1. KLASIFIKACE A PREDIKCE

Přestože cysty byly klasifikovány lékaři z Fakultní nemocnice Brno, počítačový systém má své vlastní metody klasifikace. Při klasifikaci počítačovým systémem dochází k překrývání množin bodů klasifikovaných objektů v příznakovém prostoru. Je nutno stanovit optimální práh tak, aby co



Obr. 3: Klasifikační hranice rozdělující vzorky do dvou tříd T1 a T2

nejméně cyst bylo klasifikováno špatně. Ve fázi trénování modelu klasifikace je klíčem klasifikace nalézt vhodné statistické parametry vstupních dat, které by vstupní data co nejlépe popisovala, a následně je vybrán matematický model, který přijímá tyto statistiky pro výpočty. Jeho výsledkem je stanovený práh (hranice) rozdělující vzorky do tříd.

Klasifikátory poskytujícími obvykle dobré výsledky jsou rozhodovací stromy, neuronové sítě, metody nejbližších sousedů a bayesiánské klasifikátory. V tomto projektu byl jako klasifikátor použit model bayesiánské sítě, jehož matematický model využívá podmíněné pravděpodobnosti a Bayesova pravidla. Přesnost vyhodnocení použitých cyst tohoto klasifikátoru je vysoká, dosahuje 93%.

Po získání konkrétních hodnot statistických parametrů dat je možno tyto parametry využít při predikci kategorie cysty. Cysta neznámé kategorie je podrobena měření stejných parametrů a ty jsou vloženy na vstup matematického modelu klasifikátoru. Výstupem je potestová pravděpodobnost pro danou třídu. Vzorek je zařazen do té třídy, která má nejvyšší potestovou pravděpodobnost. Pokud by se vzorek prokázal být vhodný pro zlepšení klasifikačního modelu z hlediska predikce dalšího vzorku, bude zařazen do množiny trénovacích dat mezi původní data.

## 6. ZÁVĚR

Cílem projektu bylo vytvoření doplňku k programu ImageJ, který by dokázal využít funkce programu RapidMiner pro klasifikaci a predikci třídy čelistní cysty. Tento doplněk může být dále využíván pro účely řešení podobné problematiky.

## REFERENCE

- [1] FRIEDRICH PASTER, A., VISSER, H., Stomatologická radiologie – kapesní atlas, str. 238-253, Praha: Vědecká redakce a nakladatelství Grada Publishing, a.s., 2007, ISBN: 978-80-247-1307-6.
- [2] RAPID-I, GmbH, RapidMiner brochure, Dortmund, Germany, 2001-2009 – [cit. 11.2.2013]. Dostupné na [www: http://sourceforge.net/projects/rapidminer/files/1.%20RapidMiner/4.4/](http://sourceforge.net/projects/rapidminer/files/1.%20RapidMiner/4.4/).