

Univerzita Jana Evangelisty Purkyně  
v Ústí nad Labem

Přírodovědecká fakulta



Vytváření bounding boxů ve snímcích buněk  
pořízených optickým mikroskopem

BAKALÁŘSKÁ PRÁCE

**Vypracoval:** Petr Kotlan

**Vedoucí práce:** doc. RNDr. Zbyšek Posel, Ph.D.

**Studijní program:** Matematika ve firmách a veřejné správě

Ústí nad Labem 2025



# Podklad pro zadání BAKALÁŘSKÉ práce studenta

Jméno a příjmení: Petr KOTLAN  
Osobní číslo: F21060

Téma práce: Vytváření bounding boxů ve snímcích buněk pořízených optickým mikroskopem  
Téma práce anglicky: Bounding boxes for cells in images from optical microscope  
Jazyk práce: Čeština

Související osoby: doc. RNDr. Zbyšek Posel, Ph.D. (Vedoucí)  
Katedra informatiky

## Zásady pro vypracování:

Cílem této bakalářské práce je navrhnout a implementovat metody pro automatizované vytváření bounding boxů buněk ze snímků bronchoalveolární laváže (BAL) tak, aby bylo možné generovat specifické datové sady snímků individuálních buněk. Ty mohou být v budoucnu využity pro automatickou klasifikaci buněk z BAL do jednotlivých skupin např. pomocí konvolučních neuronových sítí. Tvorba bounding boxů bude využívat segmentaci obalů buněk například pomocí tzv. Watershed algoritmu nebo pomocí techniky podpůrných vektorů (SVM). Snímky buněk z BAL budou poskytnuty na základě předchozí spolupráce s lékaři z Masarykovy nemocnice v Ústí nad Labem, o.z. Pro posouzení kvality ohraničení buněk budou v práci navrženy vhodné metriky. Vznik takovýchto datových sad, a tím i aplikace metod pro klasifikaci buněk přispěje ke zvýšení automatizace vyšetření založených na BAL snímcích, kde je cílem získat diferenciální počet jednotlivých typů buněk.

## Struktura práce:

1. Úvod do problematiky analýzy snímků buněk
2. Prezentace metod tvorby bounding boxů
3. Popis vstupních dat
4. Návrh a implementace schémat tvorby bounding boxů
5. Vyhodnocení přesnosti a vhodnosti vytvořených hranic
6. Diskuze dosažených výsledků

## Seznam doporučené literatury:

- CHECK, Irene J.; GOWITT, Gerald T. a STATON, Gerald W. Bronchoalveolar Lavage Cell Differential in the Diagnosis of Sarcoid Interstitial Lung Disease: Likelihood Ratios Based on Computerized Data Base. Online. *American Journal of Clinical Pathology*. 1985, roč. 84, č. 6, s. 744-747. ISSN 1943-7722. Dostupné z: <https://doi.org/10.1093/ajcp/84.6.744>
- FREEMAN, H. a SHAPIRA, R. Determining the minimum-area enclosing rectangle for an arbitrary closed curve. Online. *Communications of the ACM*. 1975, roč. 18, č. 7, s. 409-413. ISSN 0001-0782. Dostupné z: <https://doi.org/10.1145/360881.360919>
- KORNILOV, Anton S. a SAFONOV, Ilia V. An Overview of Watershed Algorithm Implementations in Open Source Libraries. Online. *Journal of Imaging*. 2018, roč. 4, č. 10. ISSN 2313-433X. Dostupné z: <https://doi.org/10.3390/jimaging4100123>
- MOHAMMED, Emad A.; FAR, Behrouz H.; MOHAMED, Mostafa M.A. a NAUGLER, Christopher. Application of Support Vector Machine and k-means clustering algorithms for robust chronic lymphocytic leukemia color cell segmentation. Online. In: *2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013)*. IEEE, 2013, s. 622-626. ISBN 978-1-4673-5801-9. Dostupné z: <https://doi.org/10.1109/HealthCom.2013.6720751>

Podpis studenta:

Datum:

Podpis vedoucího práce:

Datum:

## Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně a použil jen pramenů, které cituji a uvádím v přiloženém seznamu literatury.

Byl jsem seznámen s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., ve znění zákona č. 81/2005 Sb., autorský zákon, zejména se skutečností, že Univerzita Jana Evangelisty Purkyně v Ústí nad Labem má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Jana Evangelisty Purkyně v Ústí nad Labem oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

V Ústí nad Labem dne .....

Podpis: .....









VYTVÁŘENÍ BOUNDING BOXŮ VE SNÍMCÍCH BUNĚK POŘÍZENÝCH OPTICKÝM MIKROSKOPEM

**Abstrakt**

**Klíčová slova**

BOUNDING BOXES FOR CELLS IN IMAGES FROM OPTICAL MICROSCOPE

**Abstract**

**Keywords**



# Obsah

<b>Úvod</b>	<b>11</b>
<b>1 Metody předzpracování a segmentace obrazu</b>	<b>15</b>
1.1 Unsharp Masking . . . . .	15
1.2 Binarizace . . . . .	15
1.2.1 Otsu thresholding . . . . .	16
1.3 Morfologické operace . . . . .	16
1.3.1 Eroze . . . . .	16
1.3.2 Dilatace . . . . .	16
1.3.3 Otevření a uzavření . . . . .	16
1.4 Watershed . . . . .	16
1.5 Support Vector Machine . . . . .	17
1.6 Vyhodnocovací metriky . . . . .	17
1.6.1 Diceův-Sørensenův index . . . . .	17
1.6.2 Jaccardův index . . . . .	17
<b>2 Porovnání použitých metod</b>	<b>18</b>
<b>3 Segmentace snímků</b>	<b>19</b>
3.1 Segmentace jader . . . . .	19
3.2 Segmentace cytoplasmy . . . . .	19
<b>4 Tvorba Bounding Boxů</b>	<b>20</b>
<b>Seznam zdrojů</b>	<b>20</b>



# Úvod

Bronchoalveolární laváž (BAL) je invazivní vyšetření plic, které se provádí za účelem diagnostiky a monitorování plicních onemocnění. Během tohoto zákroku je do plicních sklípků zavedena tekutina, která se následně odsaje a analyzuje (BAL fluid nebo také BALF, česky BAT). Z této tekutiny je cílem získat diferenciální počet buněk jednotlivých typů buněk (příklad výstupu lze vidět na obrázku 1). Automatizovaná analýza pomocí počítačového zpracování obrazu může zrychlit diagnostiku a zlepšit přesnost.

*Table 1. Diagnoses and Bronchoalveolar Lavage Differential Cell Counts of 208 Patients Evaluated for Chronic Interstitial Lung Disease*

Diagnosis	No. Patients	Lymphocytes (%)		Neutrophils (%)		Macrophages (%)	
		Median	Range	Median	Range	Median	Range
Sarcoidosis	95	30	0–84	13	0–87	61	9–97
Idiopathic pulmonary fibrosis	35	7	0–42	12	0–63	75	0–95
Collagen vascular disease*	20	16	0–62	2	0–95	60	4–10
Pulmonary tuberculosis†	9	13	0–57	60	1–94	31	3–72
Drug-induced lung disease‡	7	14	5–65	2	0–61	61	4–90
Pneumonia, viral and/or bacterial	6	4	0–64	29	1–82	36	15–95
Primary lung cancer	5	15	4–39	7	1–80	77	5–87
Hypersensitivity pneumonitis	4	50	18–66	8	1–14	39	22–7
Pneumoconiosis	3		2–10		1–67		26–89
Other§	24	4	0–27	3	0–96	78	0–97

\* The most common etiologies were rheumatoid lung disease, systemic lupus erythematosus, and scleroderma.

† Five patients were infected with atypical Mycobacteria.

‡ The most common drugs were bleomycin and nitrofurantoin.

§ Included were a wide variety of unrelated etiologies, usually single cases such as chronic aspiration pneumonitis, ciliary motility disorder, bronchiectasis, granulomatous angitis, etc.

Obrázek 1: Diagnostika podle diferenciálního počtu buněk. Převzato z [4].

V oblasti segmentace buněk a automatizované analýzy cytologických snímků bylo provedeno několik studií zaměřených na různé metody a přístupy. Například v práci [?] autoři analyzovali využití hlubokého učení pro segmentaci a klasifikaci buněk v mikroskopických snímcích, přičemž využili konvoluční neuronové sítě (CNN) a metody transferového učení k dosažení vysoké přesnosti. Další výzkum [?] se zaměřil na použití U-Net architektury pro segmentaci buněk v biomedicínských obrazech, což vedlo ke značnému pokroku v oblasti automatizovaného zpracování obrazu. Specificky v oblasti bronchoalveolární laváže existují studie, které se snaží o automatizaci diferenciálního počtu buněk pomocí strojového učení a zpracování obrazu [?], kde byly testovány různé přístupy k detekci a klasifikaci buněk v BALF vzorcích. Tyto práce ukazují, že kombinace pokročilých metod zpracování obrazu a hlubokého učení může významně přispět k automatizaci diagnostiky v cytologii a patologii.

Základem zpracování obrazu je segmentace, jejímž cílem je detekovat a rozlišit od sebe jednotlivé buňky a jádra, případně odstranit objekty, které nejsou předmětem zkoumání. Tento úkol je v oblasti analýzy snímků buněk často obtížný. Během zpracování je třeba se vypořádat s řadou problémů, jako jsou například překrývající se buňky, šum a v neposlední řadě i samotná kvalita snímků. Buňky samotné mohou mít různé tvary a velikost. V této práci jsou v rámci oblastí segmentace jader a cytoplasmy buněk použity a porovnány algoritmy Watershed a SVM.

Hlavním cílem této práce je v segmentovaném snímku kolem buněk vytvořit bounding boxy tak, aby z jednotlivých buněk mohla být vygenerována datová sada. Tato sada bude následně použita například pro trénování konvolučních neuronových sítí, které se využívají pro klasifikaci buněk.

Práce je členěna na teoretickou část, ve které jsou popsány použité metody pro předzpracování, následnou segmentaci obrazu a vybrané metriky pro vyhodnocení kvality segmentace. V následující praktické části je nejdříve věnována kapitola porovnání dvou použitých metod pro segmentaci buněk. Následně je popsána implementace a výsledky detekce buněk a jader pomocí algoritmu watershed. V závěrečné kapitole je popsán proces vytváření bounding boxů kolem buněk.

# Metody předzpracování a segmentace obrazu

V této kapitole jsou podrobněji popsány vybrané metody předzpracování a segmentaci obrazu. Na závěr kapitoly jsou uvedeny nejpoužívanější vyhodnocovací metriky.

## 1.1 Unsharp Masking

Unsharp masking je metoda zpracování obrazu, která se používá zaostření obrázku.

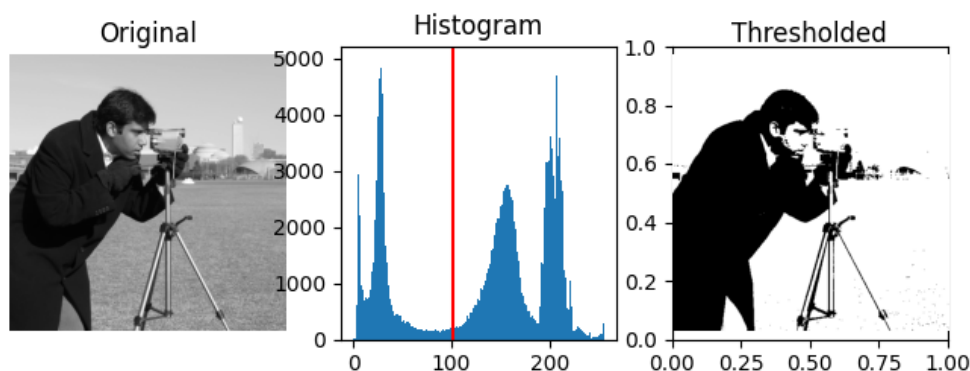
$$I_{\text{sharp}} = I_{\text{original}} + k \cdot (I_{\text{original}} - I_{\text{blurred}}) \quad (1.1)$$

$I_{\text{original}}$  je původní obrázek,  $I_{\text{blurred}}$  je rozmazaný obrázek. Nejčastěji se k rozmazání používá Gaussovský filtr.  $k$  je koeficient, který určuje sílu efektu [3].

Byla použita implementace z knihovny PIL [citace].

## 1.2 Binarizace

Metody binarizace (nebo také thresholdingu) mají za úkol nalézt prahovou hodnotu, podle které se obraz jednoznačně rozdělí na dvě třídy. Tím se oddělí pozadí od popředí.



Obrázek 1.1: Binarizace obrazu Otsu metodou. Převzato z [2].

### 1.2.1 Otsu thresholding

Metoda Otsu patří mezi nejpoužívanější metody binarizace. Algoritmus minimalizuje rozptyl uvnitř jednotlivých tříd.

$$\sigma_w^2(t) = w_0(t)\sigma_0^2(t) + w_1(t)\sigma_1^2(t) \quad (1.2)$$

$w_0(t)$  a  $w_1(t)$  jsou pravděpodobnosti přiřazení pixelu do popředí resp. pozadí,  $\sigma_0^2(t)$  a  $\sigma_1^2(t)$  jsou rozptyly tříd [1].

Metoda je implementována v knihovně Scikit-image [2].

## 1.3 Morfologické operace

Morfologické operace jsou základními nástroji pro zpracování a analýzu obrazů, které se používají k úpravě geometrických struktur v obraze. Řadí s k naprostému základu zpracování obrazu a proto zde uvedu pouze základní popis.

### 1.3.1 Eroze

Eroze je operace, která zmenšuje objekty v obraze. Používá se k odstranění malých objektů nebo šumu.

### 1.3.2 Dilatace

Dilatace je operace, která zvětšuje objekty v obraze. Používá se k vyplnění mezer nebo k propojení blízkých objektů.

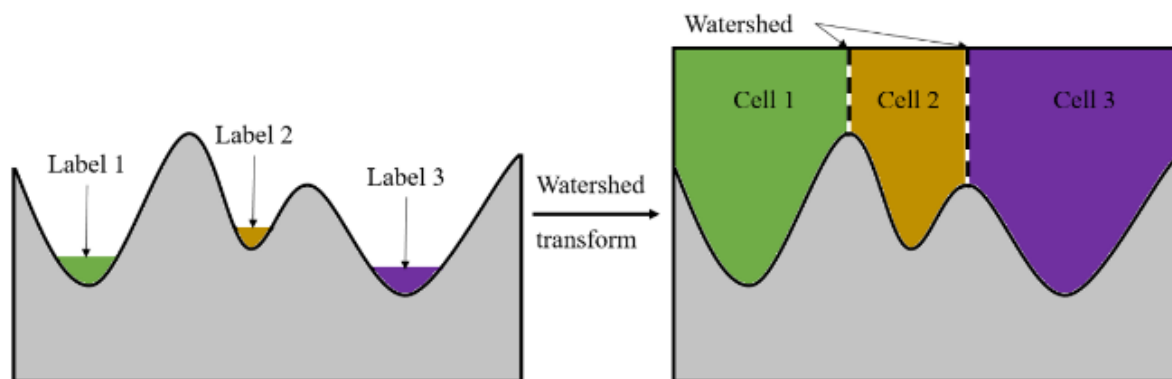
### 1.3.3 Otevření a uzavření

Otevření je kombinace eroze a dilatace, která se používá k odstranění malých objektů a šumu. Uzavření je kombinace dilatace a eroze, která se používá k vyplnění mezer a propojení blízkých objektů.

## 1.4 Watershed

Algoritmus Watershed se používá k segmentaci objektů, které se dotýkají nebo překrývají. Začne se umístěním markerů. Z každého takového bodu se pak obraz postupně „zaplavuje“, dokud se nesetkají dvě oblasti, které byly zaplavovány z různých markerů. Nejčastěji se markery určí pomocí vzdálenostní transformace, kde se určí vzdálenost každého pixelu od nejbližšího pozadí.





Obrázek 1.2: Segmentace obrazu pomocí algoritmu Watershed. Převzato z [citace].

## 1.5 Support Vector Machine

## 1.6 Vyhodnocovací metriky

### 1.6.1 Diceův-Sørensenův index

Nejpoužívanější vyhodnocovací metrikou je Diceův-Sørensenův koeficient, známý též jako F1 skóre.

$$DSC = \frac{2 \times |A \cap B|}{|A| + |B|}$$

### 1.6.2 Jaccardův index

$$J = \frac{|A \cap B|}{|A \cup B|}$$

## Porovnání použitých metod

# Segmentace snímků

## 3.1 Segmentace jader

Snímek buněk

## 3.2 Segmentace cytoplasmy

# Tvorba Bounding Boxů

# Seznam zdrojů

- [1] Otsu's method.
- [2] Thresholding, 2013.
- [3] Unsharp masking, 2013.
- [4] Irene J. Check, Gerald T. Gowitt, and Gerald W. Staton. Bronchoalveolar lavage cell differential in the diagnosis of sarcoid interstitial lung disease. *American Journal of Clinical Pathology*, 84(6):744–747, 1985-12-01.