



## Zadání diplomové práce

<b>Název:</b>	Měření tepové frekvence z kamery
<b>Student:</b>	Bc. Vít Šprachta
<b>Vedoucí:</b>	Ing. Jakub Novák
<b>Studijní program:</b>	Informatika
<b>Obor / specializace:</b>	Znalostní inženýrství
<b>Katedra:</b>	Katedra aplikované matematiky
<b>Platnost zadání:</b>	do konce letního semestru 2023/2024

### Pokyny pro vypracování

Cílem práce je vytvořit systém snímání a následného bezkontaktního měření tepové frekvence osoby. Celý systém by pak byl využitelný v rámci propagačních akcí FITu v Laboratoři zpracování obrazu (ImproLab).

Práce se zaměřuje na vytvoření snímací soustavy (kamera, objektiv, osvětlení) a dále na algoritmickou část založenou na metodách strojového vidění a zpracování obrazu (např. automatickou detekci obličeje osoby, převod do frekvenčního spektra atd.).

Úkoly:

- 1) Proveďte rešerši v oblasti bezkontaktního snímání tepové frekvence pomocí kamery.
- 2) Sestavte měřicí kamerový systém pro snímání uživatele.
- 3) Vytvořte metodiku měření včetně specifikace měřených parametrů.
- 4) Navrhněte a implementujte architekturu aplikace s důrazem na modulární strukturu řešení.
- 5) Navrhněte algoritmy pro detekce obličeje, detekce kůže a výpočet tepové frekvence pomocí metod zpracování obrazu.
- 6) Otestujte výsledky navržených algoritmů a zhodnoťte úspěšnost.
- 7) Vizualizujte výsledky algoritmů a představte funkční systém.





**FAKULTA  
INFORMAČNÍCH  
TECHNOLÓGIÍ  
ČVUT V PRAZE**

Diplomová práce

## Měření tepové frekvence z kamery

*Bc. Vít Šprachta*

Katedra aplikované matematiky  
Vedoucí práce: Ing. Jakub Novák

11. ledna 2024





---

## Poděkování

Rád bych poděkoval svému vedoucímu práce, Ing. Jakubu Novákovi, za jeho trpělivost během vedení této práce.



---

## Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principu při přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisů, zejména skutečnost, že České vysoké učení technické v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 citovaného zákona.

V Praze dne 11. ledna 2024

.....

České vysoké učení technické v Praze

Fakulta informačních technologií

© 2024 Vít Šprachta. Všechna práva vyhrazena.

*Tato práce vznikla jako školní dílo na Českém vysokém učení technickém v Praze, Fakultě informačních technologií. Práce je chráněna právními předpisy a mezinárodními úmluvami o právu autorském a právech souvisejících s právem autorským. K jejímu užití, s výjimkou bezúplatných zákonných licencí a nad rámec oprávnění uvedených v Prohlášení na předchozí straně, je nezbytný souhlas autora.*

### **Odkaz na tuto práci**

Šprachta, Vít. *Měření tepové frekvence z kamery*. Diplomová práce. Praha: České vysoké učení technické v Praze, Fakulta informačních technologií, 2024.

---

# Abstrakt

Práce se věnuje měření tepové frekvence pomocí běžné kamery. Proces měření je rozdělen do tří hlavních částí: detekce obličeje, segmentace kůže a výpočet tepové frekvence. Detekce obličeje využívá single shot detector neuronové sítě. Segmentace kůže může probíhat buď tradičními prahovacími metodami, nebo pomocí neuronové sítě pro tvorbu masky obličeje. Výpočet tepové frekvence je založen na odhadu signálu srdečního systolického objemu z jemných změn v odstínu pokožky, následovaný Fourierovou transformací k určení počtu úderů srdce za minutu. V rámci práce je navržen snímací systém a vytvořena konfigurovatelná aplikace. Výsledná přesnost při měření klidové tepové frekvence (50-80 BPM) dosahuje srovnatelných výsledků s referenčními zařízeními s chybou měření 5 %. Při měření vyšších tepových frekvencí (80-200 BPM) je chyba měření 10 %.

**Klíčová slova** počítačové vidění, tepová frekvence, bezkontaktní měření, fotopletysmografie, zpracování signálu, detekce obličeje, segmentace kůže, modulární aplikace, Python

# Abstract

The thesis deals with the measurement of heart rate using a conventional camera. The measurement process is divided into three main parts: face detection, skin segmentation and heart rate calculation. Face detection uses a single shot detector neural networks. Skin segmentation can be done either by traditional thresholding methods or by using a neural network to create a face mesh. Heart rate calculation is based on estimating the blood volume pulse from subtle changes in skin tone, followed by Fourier transform to determine the number of heart beats per minute. A camera system is designed and a configurable application is developed. The resulting accuracy in measuring resting heart rate (50-80 BPM) is comparable to reference devices with a measurement error of 5 %. At higher heart rates (80-200 BPM) the measurement error is 10 %.

**Keywords** computer vision, heart rate, contactless measurement, photoplethysmography, signal processing, face detection, skin segmentation, modular application, Python

---

# Obsah

Úvod	1
<b>1 Rešerše</b>	<b>3</b>
1.1 Klasické způsoby zisku tepové frekvence . . . . .	3
1.2 Výpočet tepové frekvence pomocí běžné kamery . . . . .	4
<b>2 Teoretický základ</b>	<b>7</b>
2.1 Využívané metriky . . . . .	7
2.2 Zpracování signálu . . . . .	8
<b>3 Analýza</b>	<b>11</b>
3.1 Snímací soustava a výpočetní systém . . . . .	11
3.2 Použité datasety . . . . .	13
3.3 Detekce obličeje . . . . .	16
3.4 Výběr kůže . . . . .	19
3.5 Výpočet tepové frekvence . . . . .	24
<b>4 Návrh funkčního systému</b>	<b>27</b>
4.1 Aplikační rozhraní . . . . .	27
4.2 Detekce obličeje . . . . .	29
4.3 Výběr kůže . . . . .	31
4.4 Výpočet tepové frekvence . . . . .	32
<b>5 Výsledky</b>	<b>37</b>
<b>6 Diskuse</b>	<b>41</b>
<b>Závěr</b>	<b>43</b>
<b>Bibliografie</b>	<b>45</b>





---

## Seznam obrázků

1.1	Umístění tepen, ze kterých se ručně měří puls [2], znázorněno červeně (základ obrázku převzat z [3]) . . . . .	3
1.2	Tonometr . . . . .	4
1.3	Ukázka BVP signálu (obrázek vytvořen z COHFACE dat [25]) . . . . .	5
2.1	Ukázka interpolace. Modře jsou původní vzorky signálu, červeně jsou interpolované hodnoty . . . . .	9
2.2	Ukázka aplikování Hammingova okénka. Modře je původní signál, červeně je signál s aplikovaným Hammingovým okénkem . . . . .	10
2.3	Ukázka Fourierovy transformace. Modře je funkce $1 \sin(1x) + 0.25 \sin(3x) + 1.5 \sin(10x)$ . Červeně je její komplexní spektrum . . . . .	10
3.1	Konfigurace kamery . . . . .	11
3.2	Snímací a Výpočetní systém . . . . .	12
3.3	Ukázky z Fddb datasetu (převzato z [28]) . . . . .	13
3.4	Ukázka z COHFACE datasetu (převzato z [30]) . . . . .	14
3.5	(a) Ukázka z datasetu UBFC 3.2.3 a (b) anonymizovaná ukázka z vlastního datasetu 3.2.4 . . . . .	15
3.6	Pipeline pro výběr detektorů obličeje . . . . .	17
3.7	Rámečky detektorů. (CvDnn žlutá, MediaPipe červená, reference bílá) . . . . .	17
3.8	Výsledky měření FPS obličejových detektorů. Svislá čára představuje hranici 8 FPS. . . . .	18
3.9	Výsledky měření atributů obličejových detektorů, hodnoty pro FPS jsou normovány z intervalu [0–100] do intervalu [0–1] . . . . .	19
3.10	16 Vybraných obrázků pro segmentační algoritmy, rozdělené do kategorií (a-d) . . . . .	21
3.11	Pipeline pro výběr metod výběru kůže . . . . .	22
3.12	Výsledky měření FPS metod pro výběr kůže . . . . .	23

3.13	Výsledky měření PPV a TPR metod pro výběr kůže pro kategorie (a-d). Černou čarou jsou propojeny pareto optimální hodnoty. . . .	23
3.14	Výsledky měření RMSE a std pro metody výběru kůže . . . . .	24
3.15	Zpracování RGB signálu na odhad tepové frekvence . . . . .	25
3.16	Výsledky měření RMSE a std pro metody výpočtu tepové frekvence	26
4.1	Ukázka UI aplikace. Na ukázce je vidět menu, detekovaný obličej, zobrazená maska kůže a odhad BPM. Obrázek na pozadí je z datasetu 3.2.3 . . . . .	28
4.2	Sekvenční diagram aplikace . . . . .	28
4.3	Výsledky měření RMSE a std pro vhodné konfigurace . . . . .	29
4.4	Diagram face-trackingu . . . . .	30
4.5	Výsledky měření RMSE a std při aplikování face-trackingu . . . .	31
4.6	Ilustrace rozdělení vysegmentované kůže na několik masek. Obrázek na pozadí je z datasetu 3.2.1 . . . . .	32
4.7	Výsledky měření RMSE a std při využití rozdělení nalezené masky	33
4.8	Výsledky měření RMSE pro měnící se délky časového okénka a tolerance . . . . .	34
4.9	Výsledky měření RMSE pro měnící se frekvenci obnovy výsledku .	35
5.1	Výsledky měření systému pro testovací část datasetu UBFC 3.2.3 .	38
5.2	Výsledky měření systému pro testovací část datasetu COHFACE 3.2.2. . . . .	38

---

## Seznam tabulek

3.1	Konfigurace systému v základní konfiguraci . . . . .	12
4.1	Vhodné metody pro jednotlivé komponenty výpočetního systému .	29
5.1	Závěrečné výsledky, 4 a 5 sloupec udávají procentuální počet výsledků v dané přesnosti, poslední sloupec udává hodnotu RMSE pro tepové frekvence menší, než 90 BPM . . . . .	37



---

## Seznam použitých zkratk

**BPM** Beats per minute, počet úderů srdce za minutu.

**BVP** Blood volume pulse, systolický objem.

**EKG** Elektrokardiografie.

**FPS** Frames per second, počet snímků za sekundu.

**HOG** Histogram of oriented gradients.

**IOU** Intersection over union, metrika.

**PPV** Positive predictive value, metrika.

**RMSE** Root mean square error, metrika.

**RPPG** Remote photoplethysmography, způsob výpočtu tepové frekvence pomocí běžné kamery.

**SSD** Single shot detector, typ neuronové sítě.

**std** Standardní odchylka, statistický ukazatel.

**TPR** True positive rate, metrika.



---

# Úvod

Srdeční tep je jedna ze základních vlastností lidského těla. Srdce čerpá krev po celém těle a rozváží tím důležité živiny a kyslík na všechna důležitá místa. Už v minulosti měření tepu využívali lékaři, či léčitelé, přinejmenším k určení smrti. Lidé ho vnímají při strachu nebo vyšší fyzické zátěži.

V dnešní době je měření tepu využíváno denně. Ze zdravotnictví se rozšířilo i do odvětví sportu, kde se využívá sportovních hodinek nebo hrudních pásů k měření tepové frekvence v reálném čase. Všechny dnes běžně rozšířené technologie mají společné to, že jsou kontaktní; vyžadují kontakt s měřeným člověkem, aby fungovaly.

Práce se zabývá bezkontaktním měřením tepové frekvence pomocí běžné kamery (zkráceně RPPG). V situacích, ve kterých nelze použít kontaktní technologie, nebo jejich použití způsobuje diskomfort, je možné využít běžnou kameru, která nabízí běžně dostupnou nekontaktní alternativu.

Případy užití mohou být v oblastech monitorování novorozenců, obětí těžkých popálenin, nebo morbidně obézních lidí. Další využití může být při sportovních zátěžových testech, nebo na propagačních akcích fakulty.



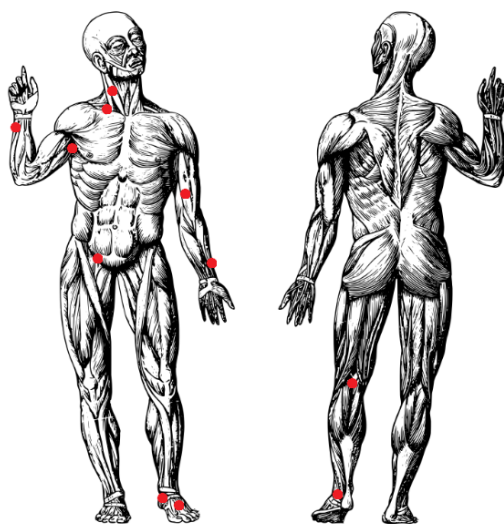


## Rešerše

V závislosti na fyzické kondici a aktivitě se srdeční tep může pohybovat v rozmezí 40 až 210 úderů za minutu (zkráceně BPM) [1]. Existují různé technologie na jeho měření.

### 1.1 Klasické způsoby zisku tepové frekvence

Základní způsob, jak někomu změřit tepovou frekvenci je přiložit prsty na jednu z tepen (obrázek 1.1) a počítat, kolik je cítit úderů srdce za časový úsek. Tento způsob se využívá zejména při lékařském fyziologickém vyšetření [2]. Jeho přesnost se odvíjí od schopnosti měřitele.



Obrázek 1.1: Umístění tepen, ze kterých se ručně měří puls [2], znázorněno červeně (základ obrázku převzat z [3])

Klinickým standardem je však měření pomocí digitálních měřičů krevního tlaku (tonometrů), které zároveň měří tepovou frekvenci pomocí oscilometrie [4, 5], nebo pulzních oxymetrů, které spolu s krevní saturací měří pulz pomocí fotopletysmografie [6]. Manuály k tonometrům uvádějí chybu měření  $\pm 5\%$  [7, 8], zatímco manuály k oxymetrům uvádějí chybu měření  $\pm 2\%$  [6, 9, 10]. Pro monitorování srdce ze závažnějších důvodů se využívá EKG [11], které dává přesné výsledky.



Obrázek 1.2: Tonometr

Oproti tomu, v oblasti sportu se používají chytré hodinky, které měří srdeční tep pomocí fotopletysmografie [12] a hrudní pásy, které pracují na základě dvoubodového EKG [13]. Chyba měření nejnovějších modelů obou technologií se pohybuje okolo 5 %, nicméně neexistuje jednotný protokol pro testování [14].

### 1.2 Výpočet tepové frekvence pomocí běžné kamery

Jedním z moderních způsobů je přímé využití neuronové sítě na vstupní video [15, 16], nicméně ostatní metody fungují na principu fotopletysmografie, s tím, že na vstupu výpočetní metody je video, ze kterého je potřeba segmentovat použitelné části kůže a až poté stanovit odhad BPM. Metodice se zkráceně říká RPPG.

Oblastí zájmu segmentace bývá čelo [17, 18, 19], celý obličej [20, 19, 21, 22, 23], nebo popřípadě jiná místa, jako ruka, nebo noha [17].

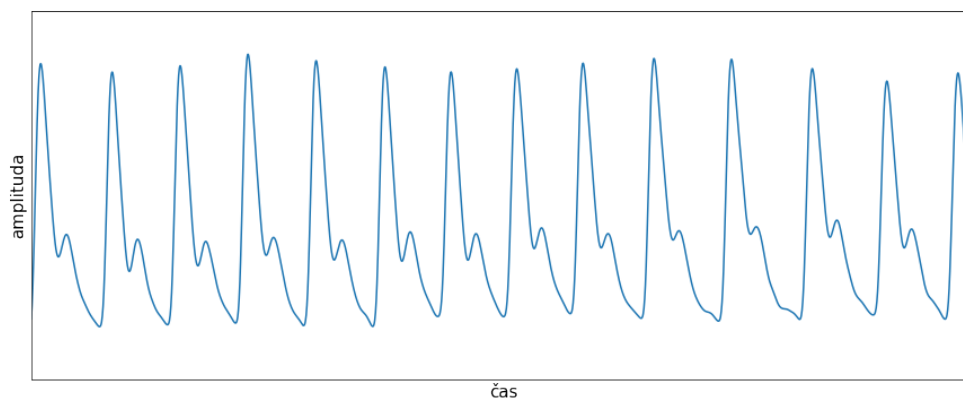
Na zvoleném místě jsou pozorovány lehké změny odstínu kůže způsobené pohybem krve, která je po těle pumpována srdcem v pulzech. V okamžiku pulzu se do žil nahrne více krve, což způsobí větší absorpci světla a nepatrné ztmavnutí kůže, které je video přijímač schopný zachytit.

Zachycením změn odstínů v čase se získá periodický signál, který slouží jako odhad systolického objemu (zkráceně BVP) [24]. Z jeho periody se následně určí výsledná tepová frekvence.

## 1.2. Výpočet tepové frekvence pomocí běžné kamery

---

Aktuální limitace RPPG spočívají v tom, že se měřený člověk nesmí hýbat a musí být dobře osvětlen. Jakýkoliv pohyb vnese šum do měřeného signálu. To se může projevit například při zvýšené fyzické aktivitě, kdy se krom tepové frekvence zvýší i dechová frekvence. Zároveň mohou být horší výsledky u lidí tmavé pleti, jelikož tmavá pleť odráží méně světla.



Obrázek 1.3: Ukázka BVP signálu (obrázek vytvořen z COHFACE dat [25])



## Teoretický základ

Kapitola poskytuje základní definice a vzorce, které jsou používány v následujících částech práce.

### 2.1 Využívané metriky

Při hodnocení účinnosti navržených metod a algoritmů jsou využívány následující metriky:

**Střední kvadratická chyba (RMSE)** Vyjadřuje míru průměrné odchylky mezi predikcemi a skutečnými hodnotami.

$$\text{RMSE} = \sqrt{\sum_{t=1}^n \frac{(\widehat{X}_t - X_t)^2}{n}} \quad (2.1)$$

**Směrodatná odchylka (std)** Udává, nakolik se od sebe navzájem typicky liší jednotlivé případy v souboru zkoumaných hodnot.

$$\sigma = \sqrt{\frac{1}{n} \sum_{t=1}^n (\widehat{X}_t - \mu)^2}, \text{ kde } \mu = \frac{1}{n} \sum_{t=1}^n \widehat{X}_t \quad (2.2)$$

**Skutečně pozitivní výsledek (TP)** Označení výsledku testu, který správně identifikoval pozitivní případ. (Pozitivní výsledek je skutečně pozitivní.)

**Falešně pozitivní výsledek (FP)** Označení výsledku testu, který nesprávně identifikoval negativní případ. (Pozitivní výsledek by měl být negativní.)

**Falešně negativní výsledek (FN)** Označení výsledku testu, který nesprávně identifikoval pozitivní případ. (Negativní výsledek by měl být pozitivní.)

**Citlivost (TPR)** Vyjadřuje pravděpodobnost pozitivního výsledku testu za předpokladu, že je jedinec skutečně pozitivní.

$$\frac{TP}{TP + FN} \quad (2.3)$$

**Pozitivní prediktivní hodnota (PPV)** Vyjadřuje pravděpodobnost pozitivního výsledku testu za předpokladu, že je výsledek testu pozitivní.

$$\frac{TP}{TP + FP} \quad (2.4)$$

**Poměr průniku a sjednocení (IOU)** Metrika vyjadřující míru překryvu mezi dvěma množinami (maskami).

$$\frac{\text{Plocha průniku A a B}}{\text{Plocha sjednocení A a B}} \quad (2.5)$$

## 2.2 Zpracování signálu

Klíčovým prvkem výpočetního systému je odhad BPM je z RGB signálu. Obecný signál můžeme definovat jako funkci  $f(t)$ , kde  $t$  značí čas a  $f(t)$  vyjadřuje hodnotu signálu v daném časovém okamžiku. RGB signál je následně specifickým typem signálu, jehož hodnotami jsou pixely.

Metodika zpracování signálu vychází z knihy [26], kde jsou detailně popsány matematické metody a postupy využívané při analýze signálů.

### 2.2.1 Interpolace

Interpolace je matematická metoda, která slouží k odhadu hodnot signálu mezi existujícími vzorky. Jejím cílem je poskytnout přesnější predikce hodnot signálu v časech, které nejsou přímo vzorkovány. V rámci práce se využívá lineární interpolace, ukázka je na obrázku 2.3.

Pokud jsou známy vzorky signálu v časech  $n = 0, 1, 2, \dots, N - 1$ , můžeme aproximovat vzorek  $f(t)$  mezi dvěma sousedními vzorky  $f(n_1)$  a  $f(n_2)$  jako

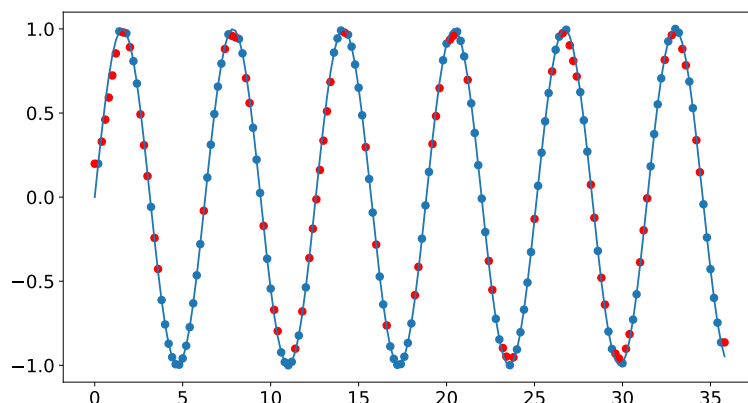
$$f(t) = f(n_1) + (t - n_1) \cdot \frac{f(n_2) - f(n_1)}{(n_2 - n_1)} \quad (2.6)$$

### 2.2.2 Hammingovo okénko

Hammingovo okénko je jedním z typů váhovacích oken používaných v digitálním zpracování signálů. Toto okénko slouží k postupnému snížení amplitudy signálu na okrajích vzorků, což pomáhá minimalizovat artefakty způsobené tzv. "únikem spektra" při Fourierově transformaci omezených časových signálů. Je definováno jako:

$$w(n) = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi n}{N-1}\right), \quad n = 0, 1, 2, \dots, N-1 \quad (2.7)$$

Kde  $w(n)$  představuje hodnotu Hammingova okénka v diskretním čase  $n$  a  $N$  je délka diskretního signálu. Grafická ilustrace Hammingova okénka je na obrázku 2.2.



Obrázek 2.1: Ukázka interpolace. Modře jsou původní vzorky signálu, červeně jsou interpolované hodnoty

### 2.2.3 Fourierova transformace

Fourierova transformace je matematická metoda, která převádí signál z časové domény do frekvenční domény. Tato transformace umožňuje reprezentovat původní signál jako lineární kombinaci sinusových a kosinusových funkcí s různými frekvencemi. Jejím hlavním cílem je odhalit frekvenční složky a jejich amplitudy obsažené v původním signálu.

Pro zpracování digitálního signálu je využívána diskretní Fourierova transformace.

$$X(k) = \sum_{n=0}^{N-1} f(n) \cdot e^{-j2\pi nk/N}, \quad k = 0, 1, 2, \dots, N-1 \quad (2.8)$$

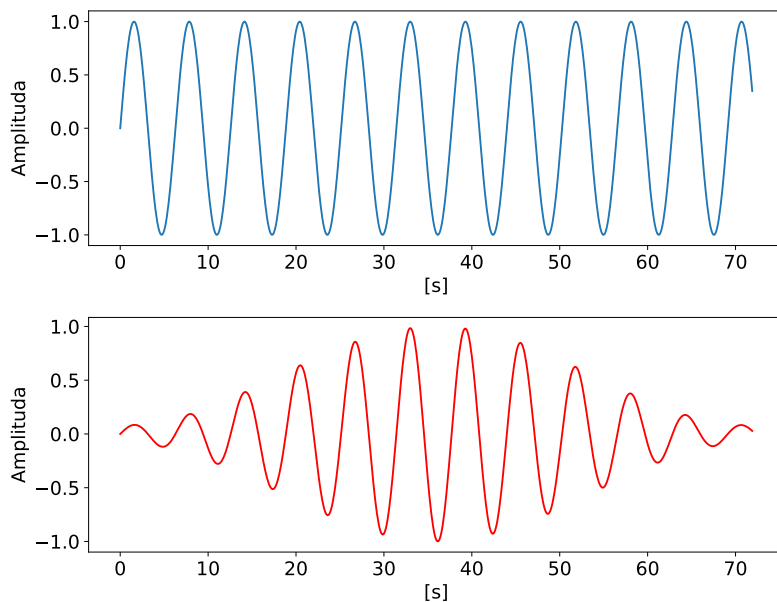
Kde  $X(k)$  je výstupní komplexní spektrum na frekvenci  $k$ ,  $f(n)$  je hodnota vstupního signálu v diskretním čase  $n$ ,  $N$  je délka diskretního signálu, a  $j$  je imaginární jednotka.

### 2.2.4 Odečtení stejnojmenné složky

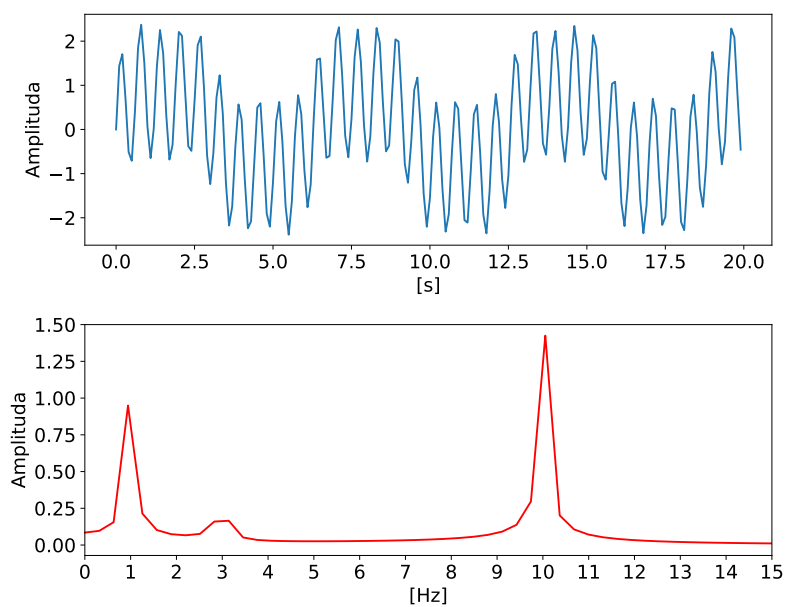
Odečtení stejnosměrné složky je proces, při kterém se ze vstupního signálu odečítá jeho střední hodnota. Tato operace slouží k centrování signálu kolem nulové hodnoty a eliminaci případných posunů v úrovni signálu.

## 2. TEORETICKÝ ZÁKLAD

---



Obrázek 2.2: Ukázka aplikování Hammingova okénka. Modře je původní signál, červeně je signál s aplikovaným Hammingovým okénkem



Obrázek 2.3: Ukázka Fourierovy transformace. Modře je funkce  $1 \sin(1x) + 0.25 \sin(3x) + 1.5 \sin(10x)$ . Červeně je její komplexní spektrum



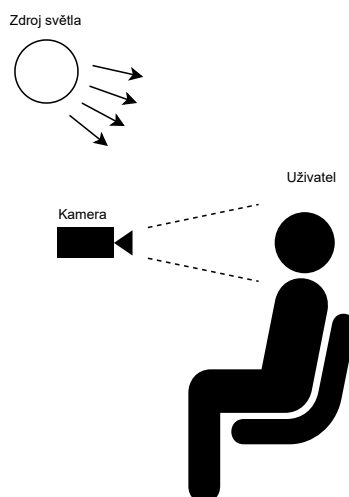
## Analýza

Kapitola poskytuje popis snímací sestavy a výpočetního systému, představuje seznam použitých datasetů pro provedení experimentů, a je v ní provedena analýza jednotlivých komponent, ze kterých je složen výpočetní systém.

### 3.1 Snímací soustava a výpočetní systém

Výsledný systém pro měření tepové frekvence z kamery je tvořen dvěma hlavními částmi: snímacím systémem a výpočetním systémem.

Snímací systém se skládá z jedné kamery, která snímá obličej uživatele zepředu. Uživatel musí mít v záběru minimálně celý obličej. Jako osvětlení postačuje přirozené světlo nebo osvětlení místnosti. Kamera musí zvládat 3 barevné složky, dosahovat alespoň 15 fps s rozlišením v rozsahu 480p až 1080p.



Obrázek 3.1: Konfigurace kamery

### 3. ANALÝZA

Výpočetní systém se skládá z tří komponent: detekce obličeje, výběr kůže a výpočet pulzu. Jejich podrobný rozbor bude proveden později v této kapitole. Systém musí zpracovávat data v reálném čase s minimální vzorkovací frekvencí 8 FPS. Například při snímání osoby s tepovou frekvencí blížící se k hornímu limitu (200 BPM) lze jeho tepovou frekvenci aproximovat jako 4 údery za sekundu.

$$200/60 = 3.\overline{33} \leq 4 \quad (3.1)$$

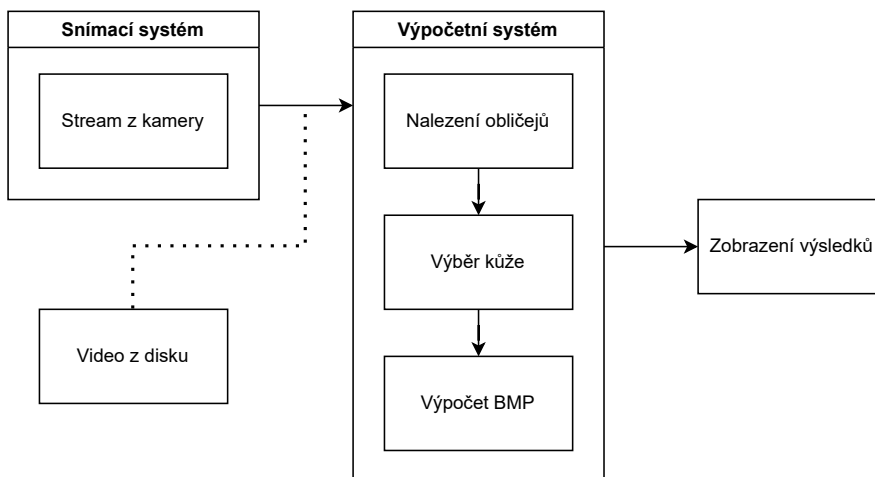
Minimální vzorkovací frekvence, nutná pro spolehlivé měření, se získá aplikací Nyquistova-Shannonova vzorkovacího teorému [26], který stanovuje minimální vzorkovací frekvenci jako dvojnásobek očekávané frekvence.

Výstupem jednotlivých měření systému je časová řada odhadů tepové frekvence. Cílem je minimalizovat chybu a variabilitu výsledků. Chyba je měřena metrikou Root Mean Square Error (RMSE) a variabilita pomocí směrodatné odchylky (std).

Veškeré experimenty jsou prováděny na notebooku vybaveném procesorem Intel Core i5-3320m @ 2.60 GHz x 4 a s 16 GB operační paměti. Termínem *základní konfigurace systému* se označuje nastavení systému podle specifikací uvedených v tabulce 3.1, přičemž detaily jednotlivých parametrů jsou popsány v následujících částech.

parametr	hodnota
Face Detector	CvDnn
Skin Selector	Forehead
Pulse Getter	raw
Časové okénko	10 s

Tabulka 3.1: Konfigurace systému v základní konfiguraci



Obrázek 3.2: Snímací a Výpočetní systém

## 3.2 Použité datasey

Pro provedení analýzy a testování systému jsou vybrány následující datasey.

### 3.2.1 FDDB dataset

FDDB dataset [27] zahrnuje přes 2000 ručně anotovaných obrázků obličejů, kde každý obrázek obsahuje jeden či více obličejů. Obrázky jsou ze široké škály prostředí, od identifikačních fotek, po momentky ze sportovních událostí.

Anotace datasetu jsou strukturovány do deseti různých částí, kde každá část obsahuje seznam obličejů. Každý obličej je popsán elipsou, která přesně definuje polohu a tvar obličeje na daném obrázku.

Dataset se primárně využívá na analýzu detekce obličejů a sekundárně při analýze segmentaci kůže.



Obrázek 3.3: Ukázky z FDDB datasetu (převzato z [28])

### 3.2.2 COHFACE dataset

COHFACE dataset [25] obsahuje 160 barevných videí, z nichž každé má délku 1 minuty, snímkovací frekvenci 20 FPS a rozlišení 640x480 pixelů. Na videích je zachyceno 40 osob sedících čelem ke kameře. K jednotlivým videím jsou zároveň přiloženy časové řady zaznamenávající systolický objem a dechovou frekvenci zaznamenaných pomocí hardwaru a softwaru od Thought Technologies<sup>1</sup>. Nicméně videa v datasetu jsou komprimovaná [29], což představuje faktor, který může ovlivnit přesnost výsledků.

<sup>1</sup><https://thoughttechnology.com/biograph-infiniti-software-platform-t7900/>

### 3. ANALÝZA

---

Osoby na videích sedí v klidu, přičemž jejich tepové frekvence jsou v rozmezí 50 až 100 BPM. Polovina videí je úmyslně osvětlena jen z jedné strany. Autoři také dataset rozdělují na trénovací a testovací část.

Dataset se využívá pro závěrečné zhodnocení přesnosti predice výpočetního systému.



Obrázek 3.4: Ukázka z COHFACE datasetu (převzato z [30])

#### 3.2.3 UBFC dataset

UBFC dataset [31] obsahuje 42 barevných videí, kde každé má délku 1 minuty, snímkovací frekvenci 30 FPS a rozlišení 640x480 pixelů. Ke každému videu je přiložen referenční záznam tepové frekvence, který byl zaznamenán pulzním oxymetrem *CMS50E*.

Měřené osoby sedí 1 metr před kamerou a hrají matematickou počítačovou hru za účelem navýšení tepové frekvence a simulování běžného interakování člověka s počítačem.

Z datasetu je vybráno 10 videí pro analýzu a testování konfigurací výpočetního systému. Testovací videa jsou vybrána tak, aby měřené osoby na videích měli tepové frekvence rovnoměrně rozdělené v rozmezí 50 až 120 BPM. Ostatní videa slouží ke zhodnocení závěrečné přesnosti systému.

#### 3.2.4 Vlastní dataset

Vlastní dataset obsahuje 9 videí juniorských sportovců po sportovní zátěži. Videá jsou dlouhá 1 minutu, v HD rozlišení, se snímkovací frekvencí 60 FPS. K videům je zároveň referenční hodnota jejich tepové frekvence ze 2 různých digitálních zařízení.

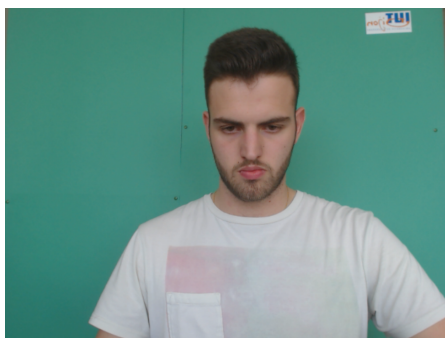
Dataset byl vytvořen na atletickém tréninku juniorských pětibojařů. Trénink probíhal venku na atletické dráze za slunného počasí s lehkou oblačností. Stanoviště pro pořizování videí do datasetu bylo sestaveno podle obrázku 3.1. Měřený dobrovolník byl měřen v sedě, na židli, natočen čelem ke kameře a slunci, které sloužilo jako přírodní osvětlení. Kamera byla umístěna na poloovatelný stativ, aby byla ve výšce obličeje, 40 cm od židle. Měřící místo bylo

umístěno tak, aby pozadí videa bylo kontrastní s kůží měřeného dobrovolníka. Jako kamera byla zvoleno *GoPro Hero 4* [32]. Referenční zařízení pro měření tepu byly pulzní oxymetr *LK87* a sportovní hodinky *Garmin Forerunner 45S*. Na základě [14] je přesnost měření obou zařízení 5 %, s tím, že měřená hodnota bude ve skutečnosti s velkou pravděpodobností vyšší.

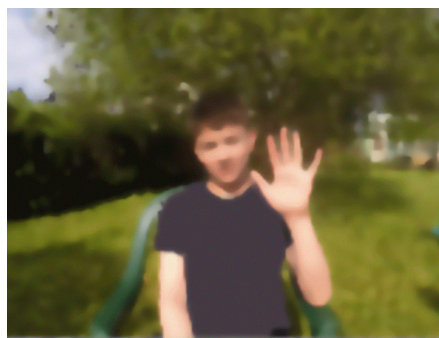
Hlavní náplní atletického tréninku byly rozběhy na 250 metrů. V každém rozběhu byl vybrán dobrovolník, který měl po čas běhu nasazené hodinky. Po doběhnutí dobrovolník usedl do židle, nasadil se mu oxymetr a po 30 sekundovém odpočinku pro zklidnění dýchání a stabilizace tepu se natočilo video. Po dokončení videa se zaznamenal srdeční tep z obou referenčních zařízení.

Na výsledných videích sedí sportovci čelem ke kameře a zároveň ukazují dlaň na kameru (více kůže v obraze). Jejich tepové frekvence se pohybují v rozmezí 90 až 150.

Dataset se využívá při vyhodnocení přesnosti predikce výpočetního systému.



(a)



(b)

Obrázek 3.5: (a) Ukázka z datasetu UBFC 3.2.3 a (b) anonymizovaná ukázka z vlastního datasetu 3.2.4

### 3.3 Detekce obličej

Vstupem pro obličejové detektory je obrázek, na němž jsou následně identifikovány obličejové a výstupem je seznam ohraničujících rámečků (bounding boxů), kde každý rámeček ohraničuje jeden nalezený obličej.

Cílem je vybrat takové obličejové detektory, které zvládnou najít obličej dostatečně rychle (alespoň 8 FPS) a nezávisle na pozadí. Rámeček, který bude následně použit pro odlišení jednotlivých lidí nebo při segmentaci kůže, by měl ideálně obsahovat celý obličej, nikoliv pouze jeho část.

Při výběru detektorů se klade důraz na následující atributy:

**FPS** Detektory musí zvládnout zpracovat alespoň 8 snímků za sekundu, aby vůbec bylo možné správně odhadnout výslednou tepovou frekvenci v rámci celého systému. Preferovány jsou detektory s vyšší hodnotou.

**IOU** Metrika vyjadřující přesnost zachycení obličejů. Měří se poměr referenční masky obličejů k nalezenému rámečku obličejů. Preferují se detektory s vyšší hodnotou.

**TPR** Metrika vyjadřující procento správně nalezených obličejů, kde TP jsou správně nalezené obličejové, a FN jsou nenalezené obličejové. Při výběru jsou preferovány detektory s vyšší hodnotou.

**PPV** Metrika vyjadřující procento správně určených obličejů, kde TP jsou správně nalezené obličejové a FP jsou nalezená místa, která nejsou obličejové. Při výběru jsou preferovány detektory s vyšší hodnotou.

#### 3.3.1 Zkoumané metody

Mezi zkoumané obličejové detektory patří jak state-of-the-art neuronové sítě, tak tradiční modely. Jejich pracovní názvy jsou **HaarCascades**, **MediaPipe**, **CvDnn**, **DlibCnn**, **MTCNN** a **RetinaFace**.

**HaarCascades** (Haarovy kaskády) jsou jeden z prvních modelů pro detekci obličejů. Z obrázku extrahují Haarovy příznaky a poté postupně použijí sérii AdaBoost klasifikátorů pro lokalizaci obličejů. Jejich výhodou je extrémní rychlost výpočtu, na druhou stranu, hlavní nevýhodou je, že nejsou invariantní vůči úhlu rotace hlavy [33].

Druhým tradičním přístupem je **HOG**. Metoda jako příznaky používá lokální gradienty obrázku, které předá do klasifikačního modelu, typicky SVM. Metoda by měla být robustnější na naklonění hlav, za cenu výpočetní síly. Problém mohou dělat tmavší obrázky, na kterých gradienty tolik nevyniknou [34, 35].

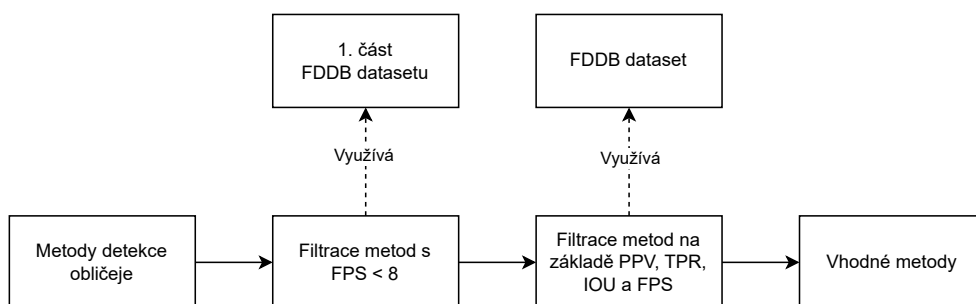
**MediaPipe** a **CvDnn** jsou zástupci tzv. Single shot detector (zkráceně SSD) neuronových sítí. Oproti ostatním neuronovým sítím přeskakují krok *návrhu oblasti zájmu* vytvořením výchozích oblastí zájmu, které jsou postupně

upravovány, aby obsahovaly chtěný objekt. Hlavní výhodou neuronových sítí je jejich celková robustnost a přesnost při predikcích. SSD sítě k tomu přidávají i rychlost výpočtu. Hlavní nevýhodou je fixní velikosti vstupu, což přidává nároky na předzpracování obrázku [36, 37].

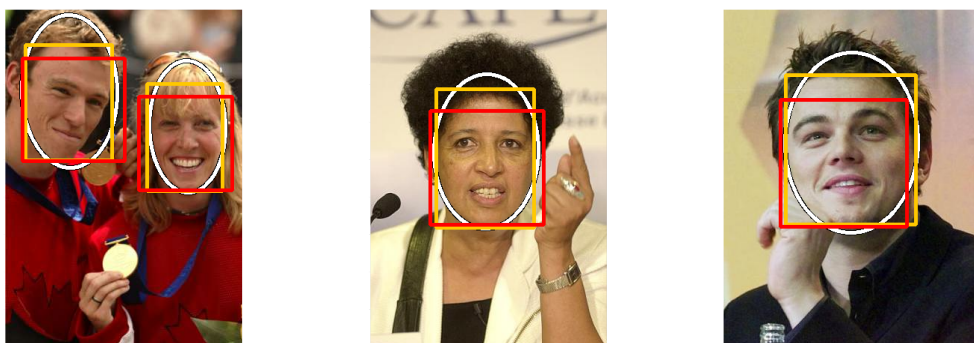
Zbylé neuronové sítě (**DlibCnn**, **MTCNN** a **RetinaFace**), leč mají rozdílné architektury, tak sdílejí stejné vlastnosti. Modely typicky vygenerují oblasti zájmu, které upraví a vyfiltrují. Jejich hlavní výhody jsou přesnost a robustnost. Hlavní nevýhodou je delší doba výpočtu [38, 39, 40, 41].

### 3.3.2 Metodika výběru

Pro změření atributů a následný výběr obličejových detektorů jsou využity obrázky z datasetu 3.2.1. Detektory jsou nejprve spuštěny pouze na první část datasetu, což umožňuje rychlé vyřazení detektorů, které nedosahují minimální rychlosti 8 FPS. Zbylé detektory jsou následně spuštěny na celý dataset, při tomto kroku jsou měřeny všechny atributy. Na základě těchto měření je provedeno zhodnocení vhodnosti jednotlivých modelů pro celkové využití v rámci systému. Celý proces je vizualizován na obrázku 3.6.



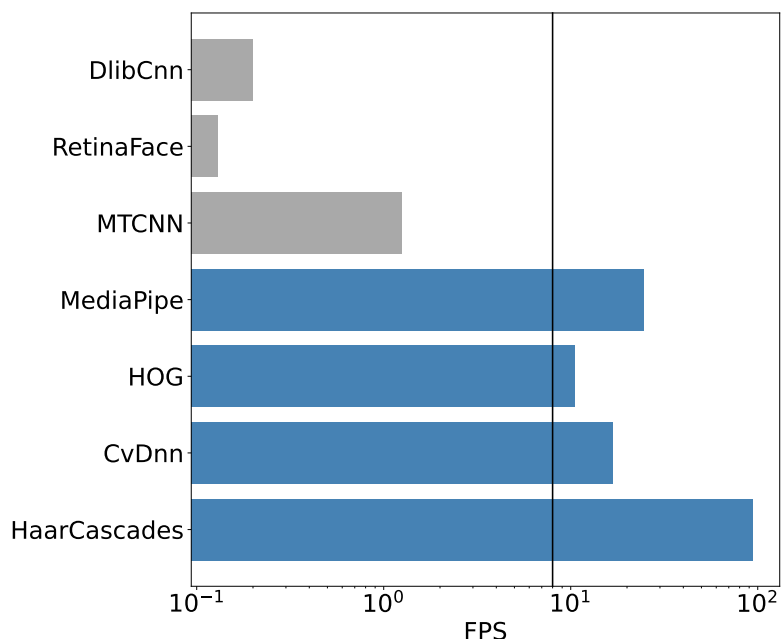
Obrázek 3.6: Pipeline pro výběr detektorů obličeje



Obrázek 3.7: Rámečky detektorů. (CvDnn žlutá, MediaPipe červená, reference bílá)

### 3.3.3 Výsledky

V první části výběru metod jsou vyřazeny detektory DlibCnn, MTCNN a RetinaFace, neboť nedosahují ani 2 FPS (viz obrázek 3.8).

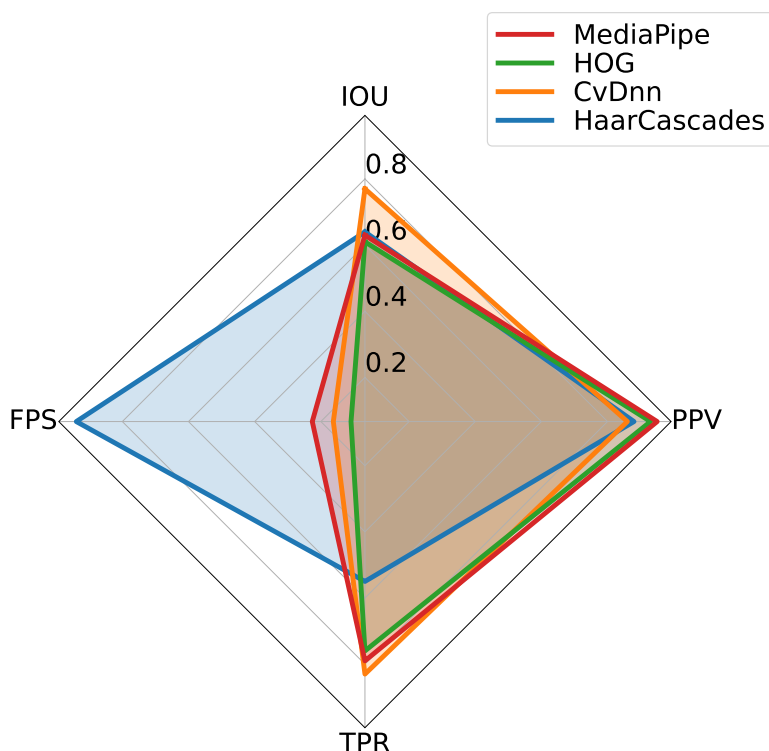


Obrázek 3.8: Výsledky měření FPS obličejových detektorů. Svislá čára představuje hranici 8 FPS.

Hodnoty naměřených atribut pro ostatní detektory jsou zachyceny na obrázku 3.9. Z výsledků je patrné, že Haarovy kaskády zabírají největší plochu v radarovém grafu, nicméně hodnota FPS nad 60 je redundantní, jelikož běžné kamery mívají snímkovací frekvenci v rozmezí 30 až 60 FPS. Zároveň hodnota TPR nedosahuje ani 60 %, což značí nedostatečnou schopnost hledání obličejů. Proto jsou Haarovy kaskády vyloučeny ze seznamu vhodných kandidátů. Spolu s nimi je vyřazen i detektor HOG, který má všechny atributy horší, než MediaPipe detektor.

Mezi zbývajícím dvěma detektory (**CvDnn** a **MediaPipe**) je zásadní rozdíl v IOU, který je vyobrazen na obrázcích 3.7. MediaPipe detektor zachycuje výrazně menší část obličeje a při manuálním testování na videu se ukazuje, že jeho velikost nalezeného rámečku není konzistentní. Obě tyto vlastnosti mohou negativně ovlivnit následný výběr kůže, proto je CvDnn vhodnějším kandidátem pro využití v rámci celého systému. Nicméně MediaPipe stále může být využit jako sekundární detektor.





Obrázek 3.9: Výsledky měření atributů obličejových detektorů, hodnoty pro FPS jsou normovány z intervalu [0–100] do intervalu [0–1]

### 3.4 Výběr kůže

Metodám pro výběr kůže na vstupu stačí samotný obrázek, nicméně přidáním rámečku obličeje se rozšíří možnosti metod o odlišení obličeje od ostatních částí těla, zúžení barevného rozsahu kůže nebo odlišení více osob na obrázku. Výstupem těchto metod je binární maska kůže, popřípadě seznam masek vysegmentované kůže. V sekci analýza se pracuje jen s kůží v obličeji.

Cílem je vybrat metody pro segmentaci kůže, které dosahují alespoň 8 FPS a dokáží vybrat kůži nezávisle na odstínu pleti a pozadí. Zároveň je kladen důraz na jejich vliv na přesnost predikce v rámci celého systému.

Při výběru metod jsou brány v potaz následující atributy:

**FPS** Selektory musí zvládnout zpracovat alespoň 8 snímků za sekundu, aby vůbec bylo možné správně odhadnout výslednou tepovou frekvenci v rámci celého systému. Preferovány jsou selektory s vyšší hodnotou.

**TPR** Metrika vyjadřující procento správně označených pixelů kůže, kde TP jsou správně detekované pixely kůže a FN jsou nedetekované pixely kůže. Při výběru jsou preferovány selektory s vyšší hodnotou.

**PPV** Metrika vyjadřující procento správně určených pixelů kůže, kde TP jsou správně detekované pixely kůže a FP jsou pixely označené jako kůže, které nejsou kůže. Při výběru jsou preferovány selektory s vyšší hodnotou.

**RMSE** Udává průměrnou chybu predikce BPM v rámci celého systému oproti referenčním hodnotám. Preferovány jsou metody s nižší hodnotou.

**std** Směrodatná odchylka udává průměrnou kvadratickou odchylku predikce BPM od jejího průměru v rámci celého systému. Preferovány jsou metody s nižší hodnotou.

#### 3.4.1 Zkoumané metody

Zkoumané metody pokrývají rozsah od jednoduchého prahování po neuronové sítě. Jejich pracovní názvy jsou **Forehead**, **HSVYCC**, **FHM**, **FloodFill**, **Watershed**, **MeanShift**, **GCut**, **FaceMesh** a **Unet**.

**Forehead** vybírá samotné čelo jako výseč rámečku obličeje. Výhoda metody je rychlost výpočtu a konzistence. Nevýhoda je, že metoda používá jen malou část dostupné kůže.

**HSVYCC** a **FHM** jsou prahovací algoritmy. První segmentuje kůži v barevných prostorech HSV a YCbCr v rámci celého obrázku na předem daných prazích pro kůži. Výhoda algoritmu je jeho rychlost, nevýhoda je, že nevyužívá informaci o nalezeném obličeji. Druhý dělá prahování celého obrázku na základě průměrné barvy čela. Algoritmus je stále rychlý a zároveň využívá nalezený obličej. Nevýhodou je, že algoritmus je náročný na dobré osvětlení obličeje [42].

**Watershed** a **FloodFill** jsou záplavové algoritmy. Jejich výhodou je využívání barevného okolí nalezeného obličeje. Nevýhodou prvního zmíněného je náchylnost k přesegmentování a nevýhodou druhého je náchylnost na změny odstínu obličeje (například vrásky) [43, 44].

**MeanShift** je shlukovací algoritmus, který shlukuje pixely do koherentních oblastí na základě jejich průměrné barvy. Jeho výhodou je, že zvládá vybrat všechnu kůži na obrázku a je invariantní vůči šumu. Hlavní nevýhodou algoritmu je jeho náročnost na výpočet [45].

**GCut** využívá GrabCut algoritmus, který odděluje pozadí od popředí pomocí prořezávání grafu vytvořeném na základě barevnosti obrázku. Výhoda algoritmu je jeho invariance vůči šumu. Nevýhoda je náročnost na výpočet a fakt, že do kůže zahrne i místa na obličeji, která kůže nejsou (například oči) [46].

**FaceMesh** je neuronová síť, která na obličeji vyhledá 478 důležitých bodů (oči, rty, tváře. . .), ze kterých následně vytvoří masku obličeje. Hlavní výhody je obecná robustnost algoritmu. Nevýhoda je, že zachycuje jen samotný obličej od brady do půlky čela [47].

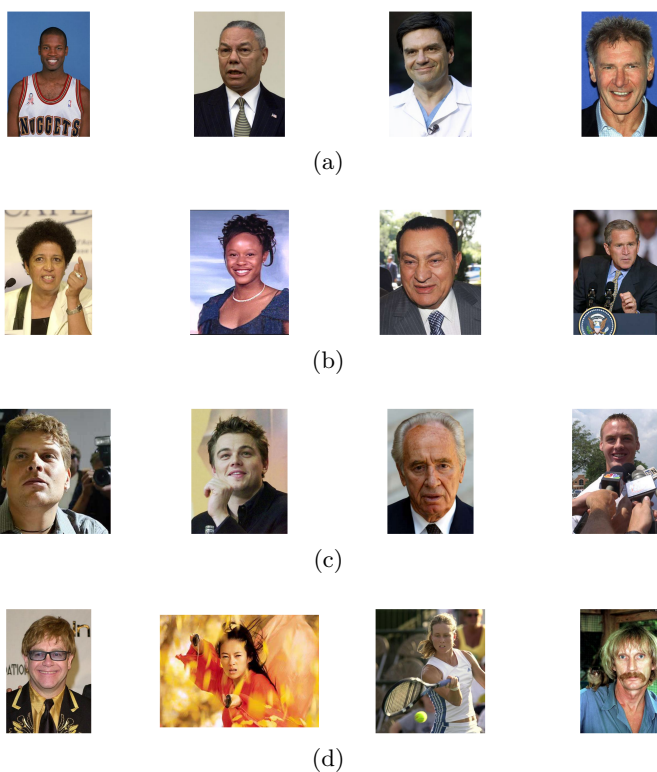
**Unet** je zástupce neuronových sítí, snaží se rovnou o segmentaci celé kůže. Výhoda metody je její robustnost. Nevýhoda je její náročnost na výpočet [48].

### 3.4.2 Metodika výběru

Pro výběr metod segmentace kůže jsou realizovány dvě fáze testování.

V první fázi jsou analyzovány atributy FPS, TPR a PPV. K tomuto účelu je vybráno 16 obrázků z datasetu 3.2.1 (viz obrázek 3.10). Tyto obrázky jsou následně rozděleny do čtyř kategorií (a-d) na základě kvality osvětlení a pozadí:

- (a) Dobré osvětlení s pozadím, které je kontrastní vůči kůži
- (b) Dobré osvětlení s různorodým pozadím
- (c) Špatné (boční) osvětlení
- (d) Fotografie s hodně rušivými prvky (dlouhé vlasy, brýle, atp.)

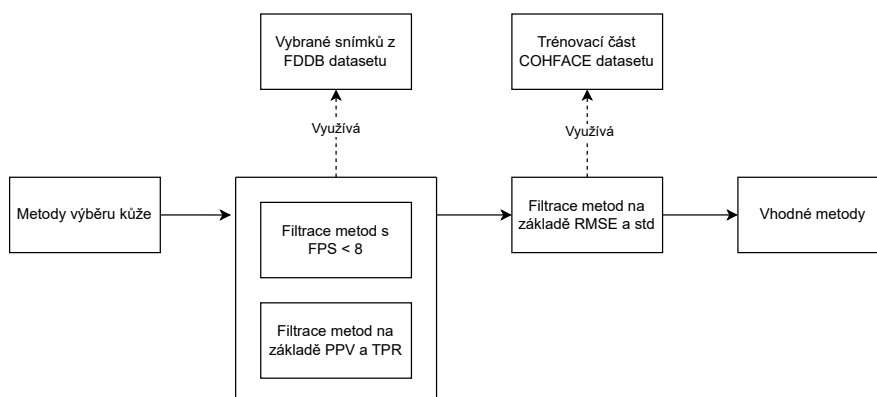


Obrázek 3.10: 16 Vybraných obrázků pro segmentační algoritmy, rozdělené do kategorií (a-d)

Každý obrázek je doplněn anotační maskou reprezentující pixely kůže.

Atributy metod jsou vyhodnoceny na základě těchto 16 obrázků. Nejprve jsou vyřazeny metody, které nedosahují 8 FPS, a následně jsou eliminovány ty, které nejsou pareto optimální [49] v kontextu PPV a TPR.

Druhá fáze testování je provedena v rámci integrace do celého systému nad trénovací částí datasetu UBFC 3.2.3. Měření jsou atributy RMSE a std. Pro tato měření je využit systém v základní konfiguraci, ve které se mění metody pro segmentaci kůže. Na základě výsledků těchto měření je vyhodnocena vhodnost metod pro praktické využití v rámci celého systému. Proces výběru je znázorněn na obrázku 3.11.



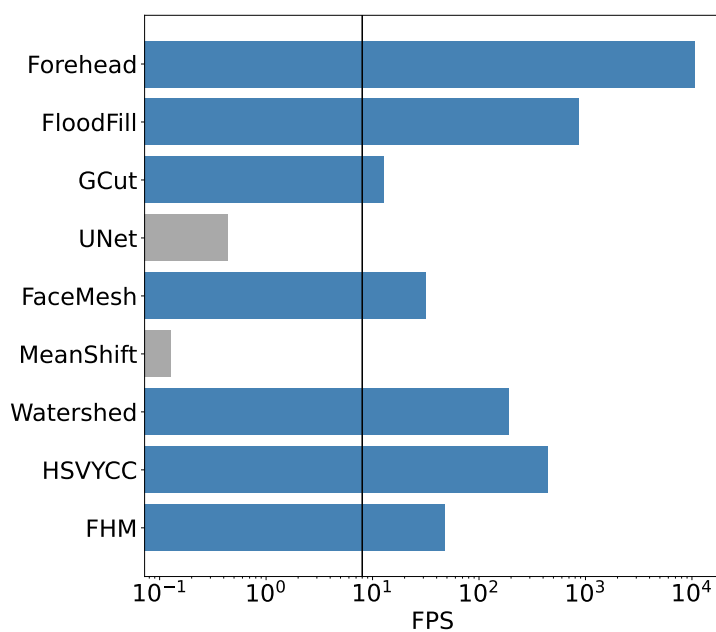
Obrázek 3.11: Pipeline pro výběr metod výběru kůže

### 3.4.3 Výsledky

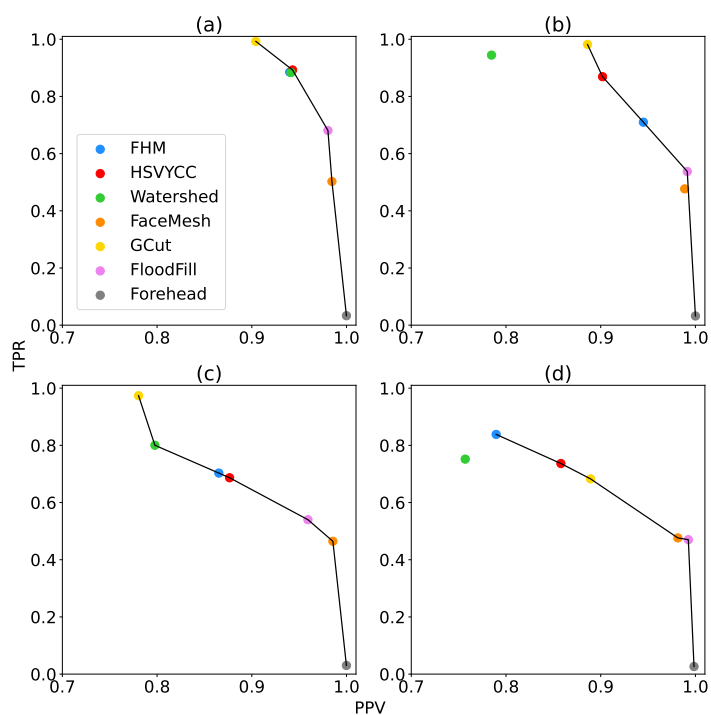
V první fázi jsou vyřazeny metody Unet a MeanShift, neboť nedosahují minimální požadované hranice 8 FPS (obrázek 3.12). Kromě nich byla také zavržena metoda Watershed, která je ve třech ze čtyř kategorií dominována jinou metodou (obrázek 3.13).

Ve druhé fázi testování vynikly dvě metody segmentace kůže: **HSVYCC** a **FaceMesh** (obrázek 3.14). Metoda HSVYCC dosahuje nejvyšší přesnosti predikce, zatímco FaceMesh exceluje v robustnosti nalezené masky a dosahuje vysoké hodnoty PPV. Obě metody jsou vyhodnoceny jako vhodné pro začlenění do celého systému.

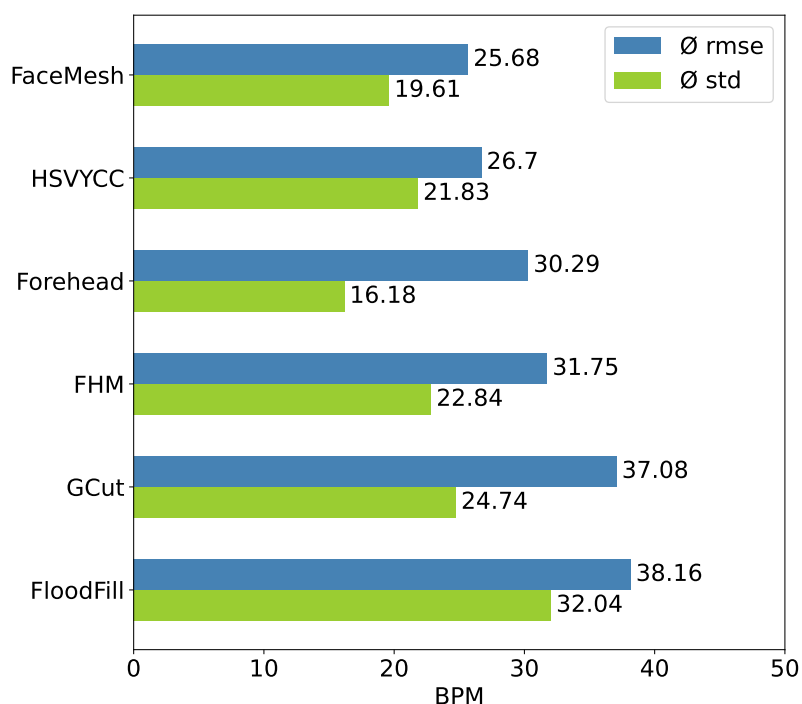
Zbývající metody nejsou klasifikovány jako vhodné. Metoda GCut má hraniční rychlost výpočtu, což je teoreticky dostačující, ale může mít v budoucnu negativní vliv na celkový výkon systému[26]. Metody FHM a FloodFill vykazují nedostatečnou konzistenci nalezených masek (viz. obrázek TODO). Nalezená maska metodou Forehead prakticky odpovídá podmnožině masky metody FaceMesh. Ač se jednoduchost a rychlost metody hodí pro základní konfiguraci systému, pro reálné nasazení převažuje výhoda rozsáhlejší a robustnější masky. Taková maska obsahuje rozsáhlejší soubor informací, což teoreticky vede k přesnějším výsledkům [50].



Obrázek 3.12: Výsledky měření FPS metod pro výběr kůže



Obrázek 3.13: Výsledky měření PPV a TPR metod pro výběr kůže pro kategorie (a-d). Černou čarou jsou propojeny pareto optimální hodnoty.



Obrázek 3.14: Výsledky měření RMSE a std pro metody výběru kůže

### 3.5 Výpočet tepové frekvence

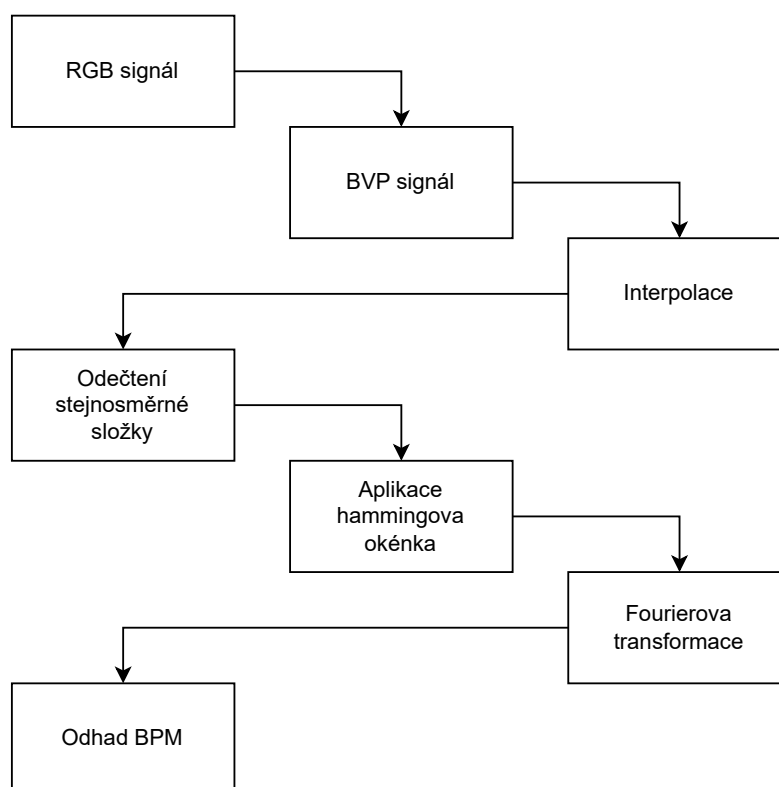
Jednotlivé metody pro výpočet tepové frekvence přijímají na vstupu RGB signál z určitého časového úseku a jejich výstupem je odhad BVP signálu. BVP signál je následně interpolován pro doplnění potenciálně chybějících hodnot. Poté je z něj odečtena stejnosměrná složka, aplikováno Hammingovo okénko a výsledný signál je podroben Fourierově transformaci k získání odhadu tepové frekvence. Celý proces je zaznamenán v obrázku 3.15. V rámci celého systému jsou tyto metody opakovaně spouštěny s každým zpracovaným snímkem videa.

Cílem této části analýzy je posoudit, které metody jsou nejvhodnější pro začlenění do celého systému, s důrazem na přesnost a stabilitu odhadu tepové frekvence. Je klíčové, aby hodnota tepové frekvence prezentovaná uživateli byla co nejpřesnější a bez výrazných výkyvů.

Měřené atributy jednotlivých metod zahrnují:

**RMSE** Udává průměrnou chybu predikce BPM v rámci celého systému oproti referenčním hodnotám. Preferovány jsou metody s nižší hodnotou.

**std** Směrodatná odchylka udává průměrnou kvadratickou odchylku predikce BPM od jejího průměru v rámci celého systému. Preferovány jsou metody s nižší hodnotou.



Obrázek 3.15: Zpracování RGB signálu na odhad tepové frekvence

### 3.5.1 Zkoumané metody

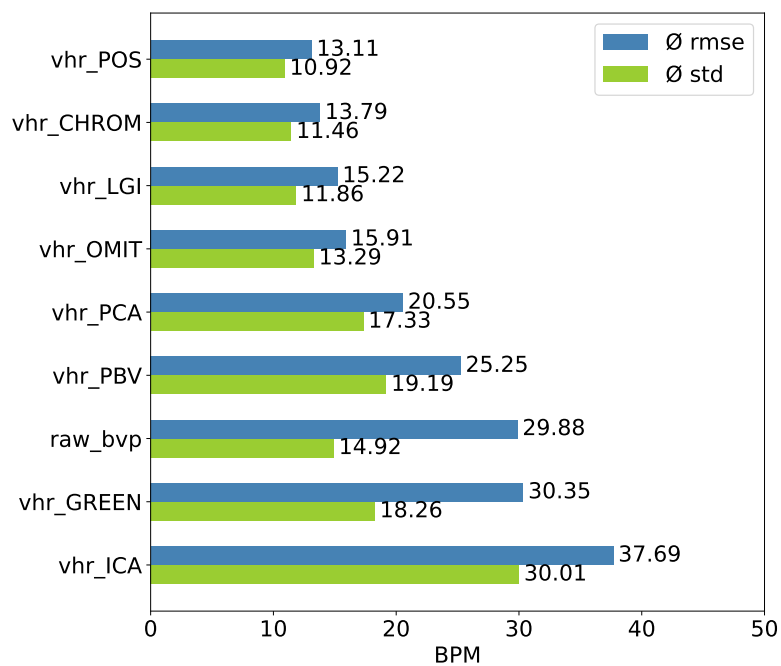
Je zkoumáno deset různých metod transformace RGB signálu na BVP signál. Podrobnější diskuse o jejich výhodách a nevýhodách je dostupná v odpovídající literatuře. Metoda **Raw** jako odhad BVP signálu používá nemodifikovaný RGB signál. Naivní přístup má výhody v rychlosti výpočtu, nicméně čistý RGB signál obsahuje i pohybový šum, který může negativně ovlivnit přesnost predikce. **CHROME** provádí normalizaci barevnosti RGB signálu s cílem odstranění pohybového šumu [51]. **GREEN** využívá pouze zelenou složku RGB signálu, jelikož obsahuje nejméně šumu [18]. **ICA** a **PCA** Používají stejnojmenné algoritmy pro oddělení BVP signálu od šumu [20, 19]. **LGI** vypočítává odhad BVP pomocí diferencovatelných lokálních transformací [52]. **POS** počítá odhad BVP z projekce RGB signálu na rovinu ortogonální na normovaný odstín kůže [21]. **OMIT** pro výpočet odhadu BVP využívá ortogonální matici získanou QR rozkladem [22]. **SSR** vychází z prostorového rozložení barev signálu a jeho časové rotace [53]. **PBV** využívá změn objemu krve v různých vlnových délkách, aby se rozlišily změny barvy způsobené pulzem od pohybových artefaktů v RGB měřeních [54].

### 3.5.2 Metodika výběru

Pro změření atributů jednotlivých metod je využita trénovací část datasetu 3.2.3. Každá z metod je integrována do celého systému v rámci jeho základní konfigurace. Je předpokládáno, že metody jsou již důkladně otestovány v předchozích pracích, a proto jsou za vhodné metody triviálně označeny tři metody s nejnižší hodnotou RMSE.

### 3.5.3 výsledky

Měření jsou vyzobrazena na obrázku 3.16. Naměřené hodnoty RMSE korelují s hodnotami std. Tři metody s nejnižšími hodnotami RMSE jsou **OMIT**, **LGI** a **CHROME**. Zároveň výsledky naznačují, že výběr správné metody pro transformaci BVP signálu je důležitější než způsob segmentace kůže.



Obrázek 3.16: Výsledky měření RMSE a std pro metody výpočtu tepové frekvence



---

# Návrh funkčního systému

Kapitola se zabývá sestavením výpočetního systému, jeho finální konfigurací a některými implementačními detaily jednotlivých modulů.

## 4.1 Aplikační rozhraní

Aplikace je implementována v jazyce Python. I přes možná omezení v efektivitě, Python přináší výhody v jednoduchosti a rychlosti vývoje, je široce používaný v oblasti počítačového vidění a pro všechny analyzované komponenty existují dostupné implementace v tomto jazyce.

Její architektura (obrázek 4.2), umožňuje snadnou konfiguraci jednotlivých komponent nebo jednoduchou integraci nových funkcí.

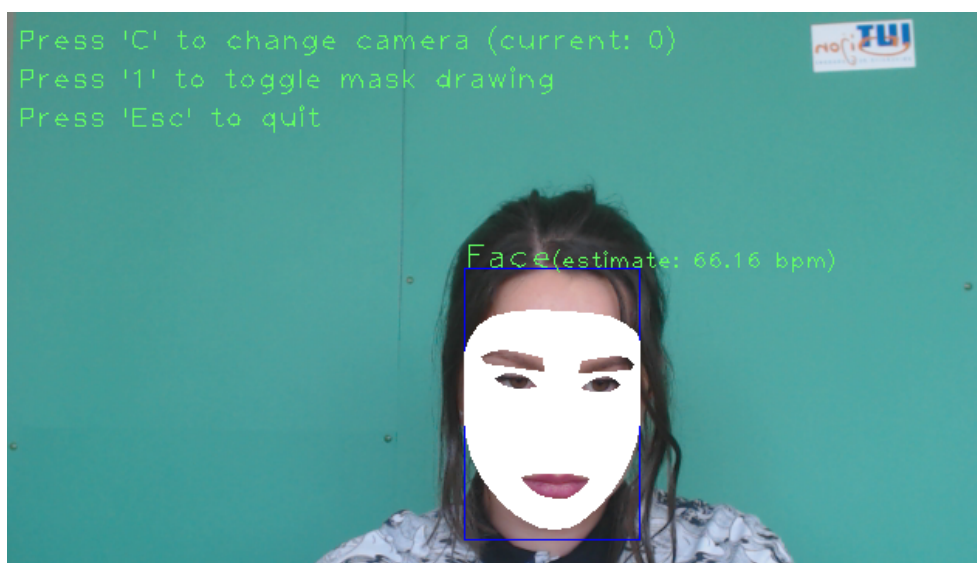
Aplikaci je možno spustit jak na real-time stream z kamery, tak na uložené video ze souboru, ke kterému se chová stejně, jako kdyby se jednalo o stream z kamery. Uživatelské rozhraní je minimalistické, ukazuje průběh videa, klávesové zkratky, aktivní obličej, odhad BPM a umožňuje zobrazit masku kůže. Počáteční konfiguraci je možno nastavit buď skrze parametry v příkazové řádce, nebo přes konfigurační *yaml* soubor.

Aktuálně je podporováno měření jedné osoby. Centrem zájmu je vždy osoba nejbližší ke středu na snímcích videozáznamu. Ukázka UI je na obrázku 4.1.

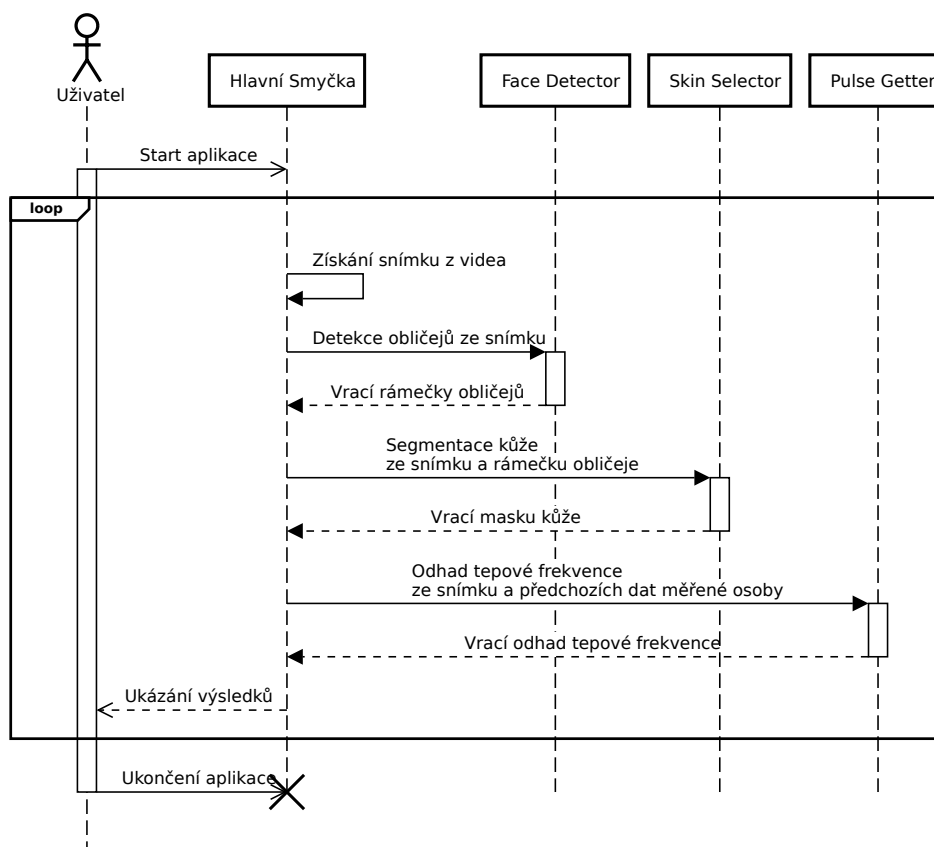
### 4.1.1 Konfigurace systému

Pro následující text a experimenty je systém vždy nastaven na konkrétní kombinace vhodných metod. Pokud není uvedeno jinak, experimenty měří přesnost odhadu (RMSE) a rozptyl odhadu (std) nad trénovací částí datasetu 3.2.3. Jednotlivé konfigurace jsou označeny třemi písmeny, kde jednotlivá písmena udávají první písmena volených způsobů pro detekci obličeje, segmentaci kůže a výpočet pulzu. Například konfigurace *CFC* značí kombinaci CvDnn, Face-*Mesh* a *CHROME*. Seznam vhodných metod je zrekapitulován v tabulce 4.1.

#### 4. NÁVRH FUNKČNÍHO SYSTÉMU



Obrázek 4.1: Ukázka UI aplikace. Na ukázce je vidět menu, detekovaný obličej, zobrazená maska kůže a odhad BPM. Obrázek na pozadí je z datasetu 3.2.3

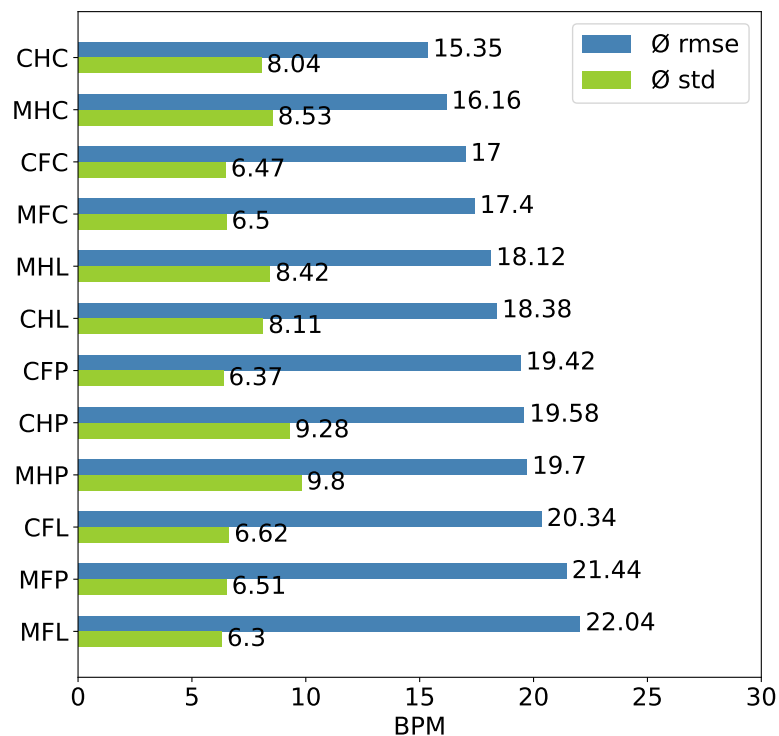


Obrázek 4.2: Sekvenční diagram aplikace

Jako baseline pro výsledky následujících experimentů je provedeno vyhodnocení přesnosti jednotlivých konfigurací. Výsledky jsou zachyceny na obrázku 4.3.

Modul	vhodné metody
Face Detector	CvDnn, MediaPipe
Skin Selector	HSVYCC, FaceMesh
Pulse Getter	CHROME, LGI, POS

Tabulka 4.1: Vhodné metody pro jednotlivé komponenty výpočetního systému



Obrázek 4.3: Výsledky měření RMSE a std pro vhodné konfigurace

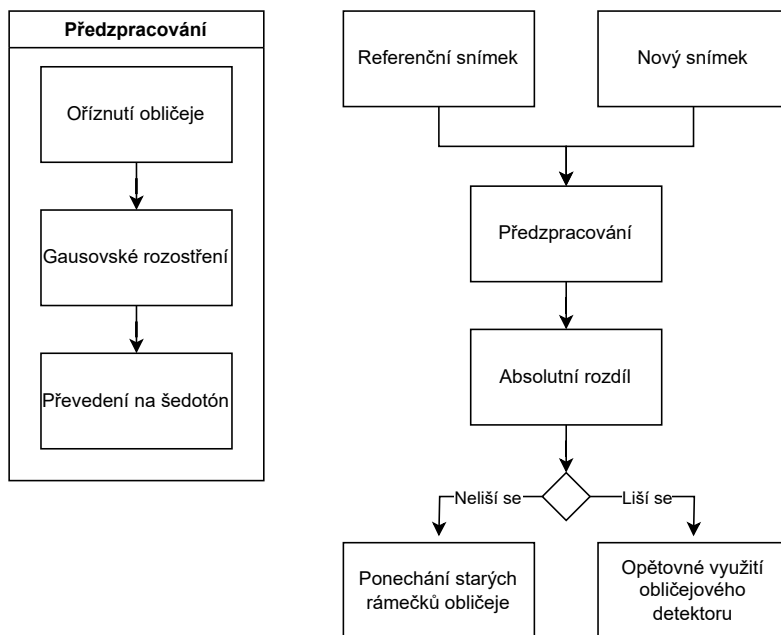
## 4.2 Detekce obličeje

Detekce obličeje představuje časově nejnáročnější část systému (viz obrázky 3.8 a 3.12), což zdůvodňuje snahu o její optimalizaci. Zvýšením počtu snímků, které systém zpracuje za sekundu lze teoreticky zlepšit výsledný odhad tepové frekvence [26]. Jednou z možností zrychlení detekce obličejů je využití sledování obličeje (face-tracking).

### 4.2.1 Face-tracking

Za předpokladu, že měřený subjekt je nehybně usazen před kameru, tak jeho obličej zůstává ve stejné pozici. Na základě tohoto předpokladu jako jednoduchý face-tracking postačuje porovnávání hodnot pixelů v oblasti obličeje mezi po sobě jdoucími snímky (obrázek 4.4).

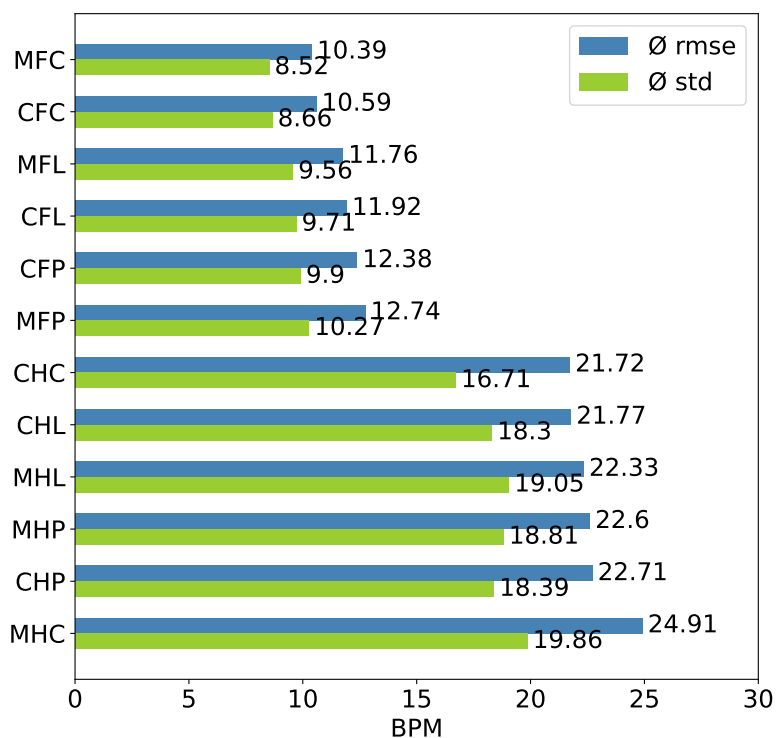
Spolu s úspěšným spuštěním detektoru je uložen aktuální snímek a nalezené obličej jako reference. Poté jsou z každého dalšího snímku extrahovány oblasti obličeje, které jsou následně převedeny na šedotón a podrobeny Gaussovskému filtru pro redukci šumu. Následně dochází k porovnání intenzit pixelů mezi aktuálním a referenčním předzpracovaným snímkem. V případě, že rozdíly nepřesahují stanovenou hranici, face-tracking je považován za úspěšný, a detektor obličeje není nutné aktivovat. V opačném případě je detektor obličeje znovu spuštěn. V každém případě se detekce obličeje aktivuje periodicky každých 5 sekund.



Obrázek 4.4: Diagram face-trackingu

### Vliv na výkon

Využití face-trackingu zvyšuje průměrné snímky za sekundu ze 12 na 20. Graf na obrázku 4.5 ukazuje, jak se tato změna projevuje na naměřené chybě. Teoreticky by naměřená chyba měla být menší, než ve výsledcích 4.3, což platí pro konfigurace, které využívají FaceMesh pro segmentaci kůže. U konfigurací využívající HSVYCC je naopak pozorováno zhoršení přesnosti. I přes to je doporučené face-tracking využívat.



Obrázek 4.5: Výsledky měření RMSE a std při aplikování face-trackingu

### 4.3 Výběr kůže

Většina používaných metod je schopna nalézt kůži i mimo obličej. Základní implemetace jednotlivých metod se však limitují jen na oblast obličeje. Výhoda tohoto omezení spočívá v redukci možných falešných pozitiv na pozadí, na druhou stranu se tím vynechávají potencionálně užitečné části kůže, jako je například krk nebo (v závislosti na použitém obličejovém detektoru) část čela.

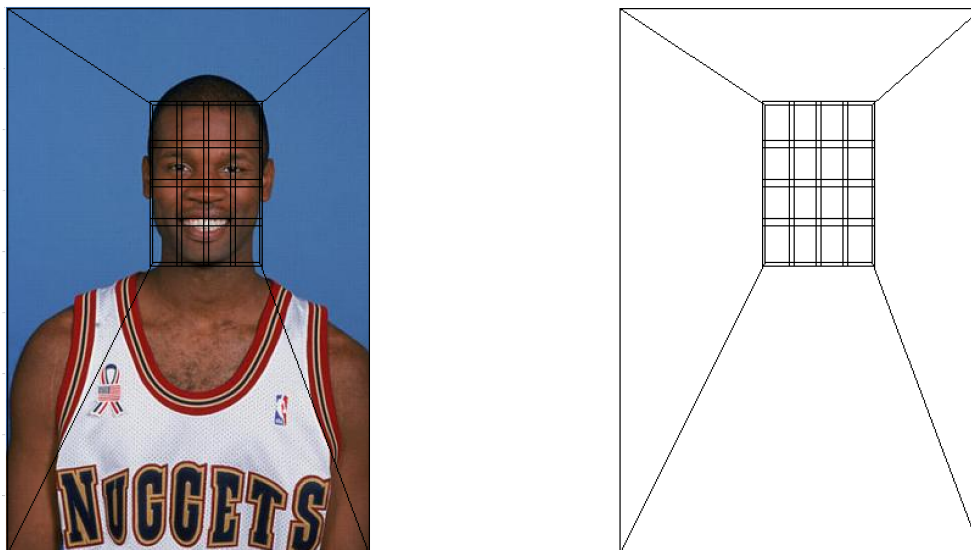
Alternativou k tomuto omezení je využití celkové vysegmentované kůže jako jedné jednotné masky. Nicméně provedené experimenty ukázaly, že tato varianta má minimální vliv na výsledky. Druhou strategií je rozdělit vysegmentovanou kůži na několik menších částí.

#### 4.3.1 Rozdělení masky na více částí

Výsledná maska je rozdělena do 20 oblastí. Čtyři oblasti jsou umístěny po stranách obličeje, zatímco obličej je dále rozdělen do 16 překrývajících se částí. Schéma tohoto rozdělení je znázorněno na obrázku 4.6.

Rozdělením obličeje na menší, překrývající se části je získáno 20 různých BVP signálů, jejichž výsledky jsou agregovány do jednoho výsledného odhadu jejich průměrnou hodnotou. Přístup se tímto snaží o zvýšení robustnosti

odhadu. Například nejednotné osvětlení (jedna strana obličeje je tmavší než druhá) nebo falešné detekce ovlivní jen jednu z částí, a tím se minimalizuje dopad na celkovou přesnost.



Obrázek 4.6: Ilustrace rozdělení vysegmentované kůže na několik masek. Obrázek na pozadí je z datasetu 3.2.1

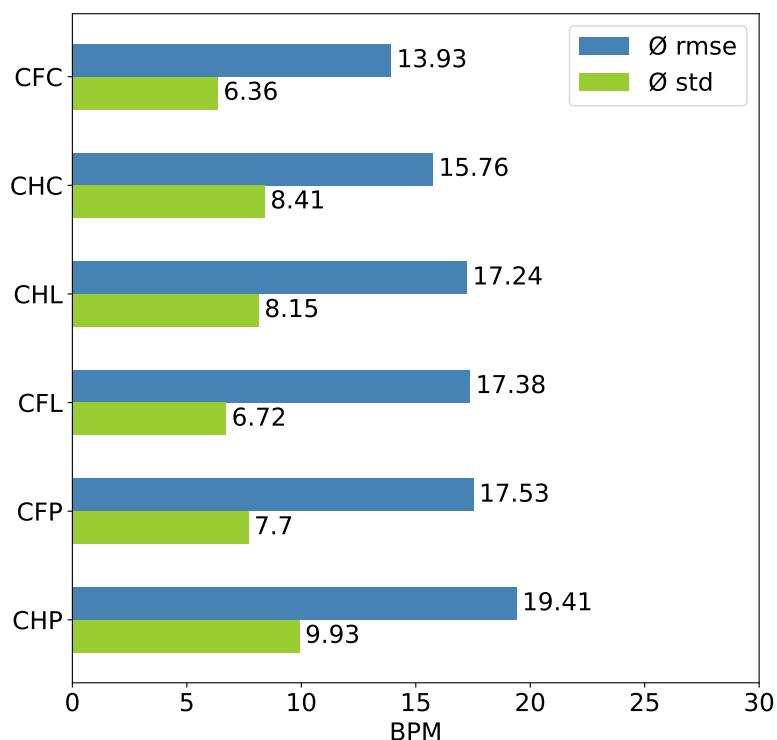
#### Vliv na výkon

U konfigurací využívajících algoritmus HSVYCC pro segmentaci kůže je pozorováno zlepšení výsledků. Naopak u ostatních konfigurací je zaznamenáno zhoršení. Toto může být způsobeno zvýšením výpočetní náročnosti, kdy systém pracuje s 20 násobkem dat. Celková rychlost systému se tímto snížila z 20 FPS (po aplikaci face-trackingu) na 16 FPS.

Z důvodu zvýšení přehlednosti a s ohledem na minimální vliv obličejového detektoru na výslednou chybu byly konfigurace zredukovány o MediaPipe detektor. Výsledky experimentu jsou zachyceny na obrázku 4.7.

### 4.4 Výpočet tepové frekvence

Metody pro odhad BVP byly navrženy na základě předchozích studií, ale stále existuje prostor pro optimalizaci a ladění systému, zejména v manipulaci s daty na vstupu a výstupu. Jde o možné úpravy délky relevantního časového okénka a frekvence obnovování výsledků.



Obrázek 4.7: Výsledky měření RMSE a std při využití rozdělení nalezené masky

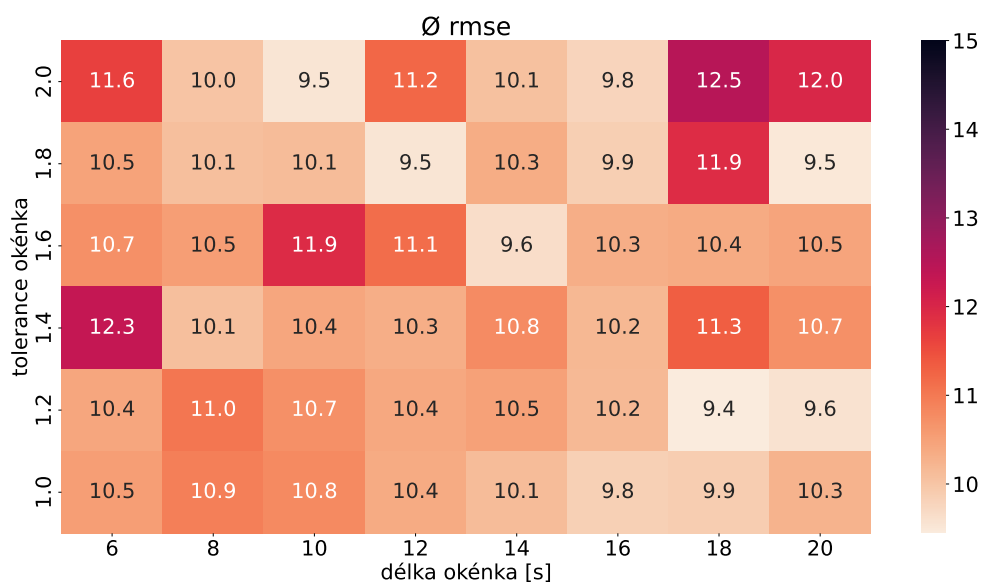
#### 4.4.1 Délka časového okénka

Základní konfigurace výpočetního systému pracuje s daty z posledních 10 sekund. Délka časového okénka je jedním z faktorů, který může ovlivnit přesnost odhadu tepové frekvence. Krátké okénko nemusí obsahovat dostatek informací pro přesný odhad, zatímco příliš dlouhé okénko může zbytečně zatěžovat paměť programu a mít problémy s odhadem u osob, jejichž tepová frekvence se rychle mění, například po fyzické námaze.

Po překročení délky časového okénka se mažou nejstarší záznamy. Protože relevantní časová data jsou uložena v poli, časté odebírání prvního prvku není efektivní. Proto je přidán druhý parametr, tolerance, který udává, kolikrát může být datový buffer delší než časové okénko, než odebere hodnoty z jeho začátku. Příliš nízká tolerance může negativně ovlivnit rychlost výpočtu, zatímco příliš vysoká může způsobit nekonzistentní výsledky kvůli razantně se měnící velikosti datového bufferu.

### Vliv na výkon

Vliv na výkon je měřen pouze pro konfiguraci CFC, která dostahuje nejpřesnějších výsledků. Výsledky (viz. obrázek 4.8) ukazují výraznou diagonálu po buňkách, kde součin hodnot délky okénka a tolerance přibližně činí 20 sekund. Na této diagonále se nacházejí optimální hodnoty. Vzhledem k tomu, že nižší hodnoty tolerance poskytují vyšší konzistenci výsledku, je ideální zvolit délku okénka 18 sekund s tolerancí 1,2.



Obrázek 4.8: Výsledky měření RMSE pro měnící se délky časového okénka a tolerance

#### 4.4.2 Frekvence obnovování výsledku

Další možná optimalizace spočívá v upravení obnovovací frekvence zobrazeného výsledku. Pokud systém dosahuje rychlosti 20 FPS, není potřeba obnovovat hodnotu zobrazenou uživateli 20 krát za sekundu. Každé obnovení znamená opětovné spuštění metody pro výpočet pulzu a s tím spojené předzpracování signálu následované Fourierovou transformací.

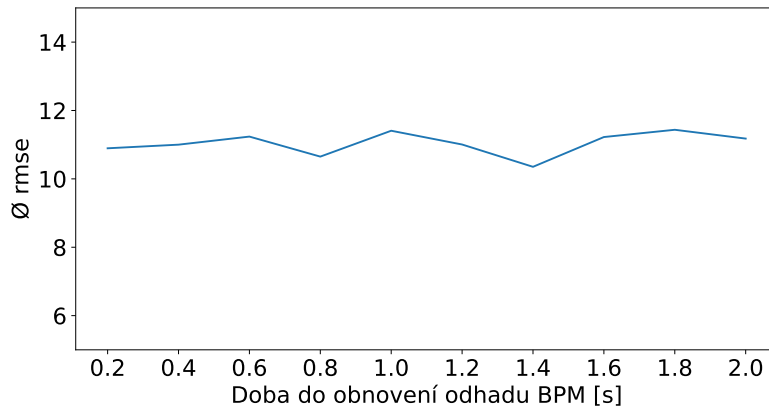
### Vliv na výkon

Výsledky (obrázek 4.9) ukazují, že optimalizace obnovovací frekvence nemá praktický vliv na přesnost výsledků. Nicméně ponechání frekvence kolem jedné až dvou obnov za sekundu může přispět k lepší uživatelské přívětivosti, kdy si uživatel stihne odhadovanou hodnotu přečíst.



#### 4.4. Výpočet tepové frekvence

---



Obrázek 4.9: Výsledky měření RMSE pro měnící se frekvenci obnovy výsledku



## Výsledky

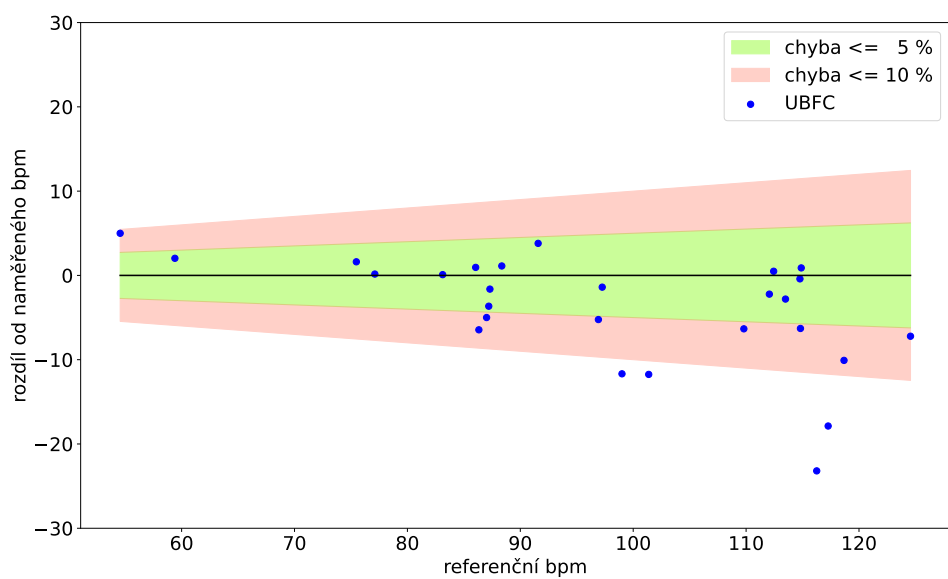
Kapitola se zavývá konečnou přesností systému. Na základě analýz je zvolena jedna nejpřesnější konfigurace systému, u které je vyhodnocena její přesnost napříč všemi datasey. Tato konfigurace se skládá z: **CvDnn** jako obličejového detektoru, **FaceMesh** metody pro segmentaci kůže a **CHROME** jako algoritmu na získání odhadu BVP. Je povolen face-tracking, časové okénko je nastaveno na 18 sekund s tolerancí 1,2. Masky kůže je brána jako jedna jednodílná maska.

Výsledky jsou prezentovány nejprve v souhrnné tabulce a následně je každému datasetu věnována samostatná část. COHFACE dataset je rozdělen na dvě části: C pro normální osvětlení a N pro podmínky špatného osvětlení (viz 3.2.2). Vlastní dataset je prezentován ve své vlastní sekci a není zahrnut v tabulce.

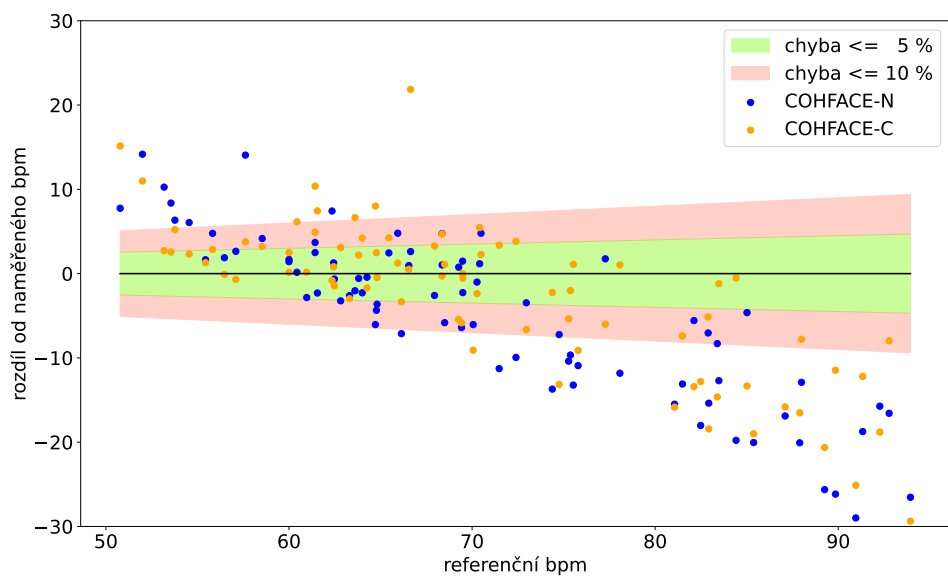
Dataset	RMSE	std	0.05	0.1	kl-RMSE
UBFC	12.8	9.6	0.54	0.82	6.4
COHFACE-C	14.3	11.8	0.39	0.68	13.7
COHFACE-N	15.8	12.6	0.35	0.57	15.0

Tabulka 5.1: Závěrečné výsledky, 4 a 5 sloupec udávají procentuální počet výsledků v dané přesnosti, poslední sloupec udává hodnotu RMSE pro tepové frekvence menší, než 90 BPM

## 5. VÝSLEDKY



Obrázek 5.1: Výsledky měření systému pro testovací část datasetu UBFC 3.2.3



Obrázek 5.2: Výsledky měření systému pro testovací část datasetu COHFACE 3.2.2.

---

## UBFC dataset

Dataset UBFC poskytuje nejlepší simulace nasazení systému do běžných podmínek. Celkovou přesnost systému zde zhoršují zejména mezní hodnoty okolo 120 BPM. Pro měření klidové tepové frekvence je systém srovnatelný s kontaktními alternativy, kdy průměrná naměřená chyba činí 6,4 BPM, což odpovídá přibližně 5-10% odchylce.

## COHFACE dataset

Výsledky naznačují, že systém je robustní vůči špatnému osvětlení, protože nejsou pozorovány zřetelné rozdíly mezi variantami datasetu. Na druhou stranu je patrné, že systém vykazuje tendenci odhadovat frekvenci 60 BPM, bez ohledu na vstupní data. Toto chování však může být zaviněno ztrátovou kompresí datasetu, která prakticky maže potřebné informace pro přesný odhad tepové frekvence. Tuto domněnku podporuje i výrazně vyšší přesnost odhadu na UBFC datasetu.

## Vlastní dataset

Výsledky měření u vlastního datasetu mají naměřené hodnoty RMSE kolem 60, tzn. systém není schopný správně odhadnout tepovou frekvenci pro videa z toho datasetu. To může být zapříčiněno několika faktory. Obecně je pozorováno zhoršení přesnosti pro vyšší tepové frekvence ve všech datasetech, avšak u tohoto datasetu je to ještě zvýrazněno tím, že snímané osoby jsou měřeny ihned po sportovní aktivitě a oddechují, což do měřeného signálu vnáší nechtěný pohyb. Druhým faktorem je vyšší rozlišení kamery, které je násobně vyšší než u předchozích datasetů. To vede k vyšším nárokům na výpočet, jak pro detekci obličeje z hlediska prohledávané oblasti, tak i pro ostatní metody, které musí pracovat s mnohem rozsáhlejším souborem dat. Důsledkem je, že snímkovací frekvence videa dosahuje pouze **8 FPS**, což se pohybuje na samotné teoretické hranici úspěšného odhadu, avšak prakticky je to evidentně nedostačující.

Zajímavým poznatkem je také, že hodnoty naměřené referenčními zařízeními mají mezi sebou odchylku 5 %. Tato odchylka odpovídá rešerši, kde každý výrobce posuzuje přesnost svého zařízení podle svých interních metrik.

Závěrem lze konstatovat, že tento experiment jednoznačně ukazuje, že vyšší rozlišení kamery nevede nutně k přesnější predikci. Naopak, zvyšování rozlišení kamery může způsobit celkové zhoršení výsledků v důsledku zvýšených nároků na výpočet.



---

## Diskuse

Pokud je systém nakonfigurován tak, že pracuje pouze s jednou jednodlitou maskou nebo využívá FaceMesh pro segmentaci kůže, lze teoreticky zvážít úplné vynechání detekce obličeje, za předpokladu, že měření má být prováděno pouze na jedné osobě. S ohledem na zrychlení, které přináší použití face-trackingu, by tato optimalizace mohla dále zvýšit přesnost systému. Jako jednu z možných implementací je možné uvažovat vracení jednoho rámečku obličeje o velikosti celé obrazovky při použití obličejeového detektoru. V případě potřeby měření více osob současně by bylo nezbytné ponechat rámečky a zároveň implementovat algoritmus pro přiřazení jednotlivých rámečku konkrétním obličejům.

Při výběru kůže dominuje metoda FaceMesh, využívající neuronovou síť. Na rozdíl od ostatních metod poskytuje konzistentní výsledky bez ohledu na podmínky. Její limitace spočívá v omezení na body v obličeji. Pro další výzkum by bylo zajímavé natrénovat podobný model neuronové sítě, který by byl schopný zachytit širší kontext lidského těla, zejména krk.

Přístup s rozdělením jedné masky na několik menších masek se neosvědčil, pravděpodobně kvůli časově náročnějšímu výpočtu. Přesto tato metoda přináší zajímavou myšlenku, kterou lze dále rozvíjet, například prostřednictvím porovnávání různých agregačních funkcí mimo jednoduché průměrování napříč maskami.

Další celkové zrychlení výpočetního systému by mohlo být dosaženo přepsáním některých částí do jiného programovacího jazyka, například do C++. Ačkoliv klíčové moduly aplikace využívají knihovny NumPy a OpenCV, které jsou implementované v C/C++, komunikace s nimi a předávání dat obstarává Python. Například časová řada, ze které se počítá odhad BVP, je uložena ve struktuře Python list a je periodicky překopírovávána do NumPy array. Tento proces, ač je neefektivnější využití Pythonu, by bylo možné použitím C++ zrychlit minimálně dvakrát.





---

## Závěr

Práce je zaměřena na bezkontaktní snímání tepové frekvence pomocí kamery. Je navržen snímací systém a jeho prototyp je využit pro tvorbu vlastního datasetu. Spolu se snímacím systémem je navržena funkční modulární aplikace se třemi hlavními komponentami: detekce obličeje, segmentace kůže a výpočet tepové frekvence. Pro každou komponentu je vypracována specifická metodika měření, na základě níž jsou vybrány adekvátní algoritmy pro danou komponentu. Výsledky měření jsou průběžně vizualizovány.

Nejllepších výsledků dosahuje konfigurace využívající neuronovou síť CvDnn pro detekci obličeje, neuronovou síť FaceMesh pro výběr kůže na obličeji a algoritmus CHROME pro výpočet tepové frekvence, který normalizuje barevé spektrum získaného signálu.

Přesnost systému pro měření klidové tepové frekvence (50-100 BPM) dosahuje systém přesnosti srovnatelné s referenčními zařízeními, s odchylkou do 5 %. Pro vyšší tepové frekvence (100-200 BPM) je naměřená chyba 10 %.



---

## Bibliografie

1. KOLEKTIV, Kittnar Otomar a. *Lékařská fyziologie*. Grada Publishing, 2011.
2. CHROBÁK, Ladislav. *Propedeutika vnitřního lékařství-Nové, zcela přepracované a doplněné vydání*. Grada Publishing, 2007.
3. J4P4N. *Human Muscle System* [online]. 2020. [cit. 2023-04-17]. Dostupné z: <https://openclipart.org/detail/326556/human-muscle-system>.
4. ALPERT, Bruce S.; QUINN, David; GALLICK, David. Oscillometric blood pressure: a review for clinicians. *Journal of the American Society of Hypertension*. 2014, roč. 8, č. 12, s. 930–938. ISSN 1933-1711. Dostupné z DOI: <https://doi.org/10.1016/j.jash.2014.08.014>.
5. LIU, Jiankun; CHENG, Hao-Min; CHEN, Chen-Huan; SUNG, Shih-Hsien; MOSLEHPOUR, Mohsen; HAHN, Jin-Oh; MUKKAMALA, Ramakrishna. Patient-specific oscillometric blood pressure measurement. *IEEE Trans. Biomed. Eng.* 2016, roč. 63, č. 6, s. 1220–1228.
6. MILNER, Q. J. W.; MATHEWS, G. R. An assessment of the accuracy of pulse oximeters. *Anaesthesia*. 2012, roč. 67, č. 4, s. 396–401. Dostupné z DOI: [10.1111/j.1365-2044.2011.07021.x](https://doi.org/10.1111/j.1365-2044.2011.07021.x).
7. OMRON. *Digital Automatic Blood Pressure Monitor Model M3: Instruction Manual* [online]. [B.r.]. [cit. 2023-04-17]. Dostupné z: <https://medaval.ie/docs/manuals/Omron-HEM-7200-E-Manual.pdf>.
8. TESLA. *Tesla Smart Blood Pressure Monitor: User Manual* [online]. [B.r.]. [cit. 2023-04-17]. Dostupné z: <https://www.teslasmart.com/data/files/268-manual-tesla-smart-blood-pressure-monitor.pdf>.
9. SILVERCREST. *Pulse Oximeter: Instruction Manual* [online]. [B.r.]. [cit. 2023-04-17]. Dostupné z: <https://manualzz.com/doc/en/57475851/silvercrest-290989-instructions-for-use-and-safety>.

10. SANTAMEDICAL. *Fingertip Pulse Oximeter: Instruction Manual* [online]. [B.r.]. [cit. 2023-04-17]. Dostupné z: <https://www.manualslib.com/manual/1396399/Santamedical-Sm-210.html>.
11. MIRVIS, David M; GOLDBERGER, Ary L. Electrocardiography. *Heart disease*. 2001, roč. 1, s. 82–128.
12. MACDONALD, Paul R.; KULACH, Christopher J. *Heart Rate Monitor With Time Varying Linear Filtering*. Vynálezce: Paul R. MACDONALD; Christopher J. KULACH. Publ.: 2017-10-31. Utility Patent US 9801587B2. Dostupné také z: <https://patents.google.com/patent/US9801587B2/>.
13. NISSILÄ, Seppo. *Method and Apparatus for Monitoring Exercise with Heart Rate*. Vynálezce: Seppo NISSILÄ. Publ.: 2004-08-10. Utility Patent US6775566B2. Dostupné také z: <https://patents.google.com/patent/US6775566B2/>.
14. COSOLI, Gloria; SPINSANTE, Susanna; SCALISE, Lorenzo. Wrist-worn and chest-strap wearable devices: Systematic review on accuracy and metrological characteristics. *Measurement*. 2020, roč. 159, s. 107789. ISSN 0263-2241. Dostupné z DOI: <https://doi.org/10.1016/j.measurement.2020.107789>.
15. LIU, Xin; FROMM, Josh; PATEL, Shwetak; MCDUFF, Daniel. Multi-Task Temporal Shift Attention Networks for On-Device Contactless Vitals Measurement. In: LAROCHELLE, H.; RANZATO, M.; HADSELL, R.; BALCAN, M.F.; LIN, H. (ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2020, sv. 33, s. 19400–19411. Dostupné také z: [https://proceedings.neurips.cc/paper\\_files/paper/2020/file/e1228be46de6a0234ac22ded31417bc7-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2020/file/e1228be46de6a0234ac22ded31417bc7-Paper.pdf).
16. SPETLIK, Radim; CECH, Jan; FRANC, Vojtěch; MATAS, Jiri. Visual Heart Rate Estimation with Convolutional Neural Network. In: 2018.
17. KOOLIJ, Koen M. van der; NABER, Marnix. An open-source remote heart rate imaging method with practical apparatus and algorithms. *Behavior Research Methods*. 2019, roč. 51, č. 5, s. 2106–2119. Dostupné z DOI: 10.3758/s13428-019-01256-8.
18. VERKRUYSSSE, Wim; SVAASAND, Lars O; NELSON, J Stuart. Remote plethysmographic imaging using ambient light. *Optics Express*. 2008, roč. 16, č. 26, s. 21434. Dostupné z DOI: 10.1364/oe.16.021434.
19. LEWANDOWSKA, Magdalena; RUMIŃSKI, Jacek; KOCEJKO, Tomasz; NOWAK, Jędrzej. Measuring pulse rate with a webcam — A non-contact method for evaluating cardiac activity. In: *2011 Federated Conference on Computer Science and Information Systems (FedCSIS)*. 2011, s. 405–410.

20. POH, Ming-Zher; MCDUFF, Daniel J.; PICARD, Rosalind W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt. Express*. 2010, roč. 18, č. 10, s. 10762–10774. Dostupné z DOI: 10.1364/OE.18.010762.
21. WANG, Wenjin; BRINKER, Albertus C. den; STUIJK, Sander; HAAN, Gerard de. Algorithmic Principles of Remote PPG. *IEEE Transactions on Biomedical Engineering*. 2017, roč. 64, č. 7, s. 1479–1491. Dostupné z DOI: 10.1109/TBME.2016.2609282.
22. CASADO, Constantino Álvarez; LÓPEZ, Miguel Bordallo. *Face2PPG: An unsupervised pipeline for blood volume pulse extraction from faces*. 2022. Dostupné z arXiv: 2202.04101 [cs.CV].
23. BOCCIGNONE, Giuseppe; CONTE, Donatello; D'AMELIO, Alessandro; CUCULO, Vittorio; GROSSI, Giuliano; LANZAROTTI, Raffaella; MORTARA, Edoardo. pyVHR: a Python framework for remote photoplethysmography. *PeerJ Computer Science*. 2022, roč. 8, e929.
24. PEPPER, Erik; HARVEY, Richard; LIN, I-Mei; TYLOVA, H.; MOSS, Donald. Is There More to Blood Volume Pulse Than Heart Rate Variability, Respiratory Sinus Arrhythmia, and Cardiorespiratory Synchrony? *Biofeedback*. 2007, roč. 35, s. 54–61.
25. HEUSCH, Guillaume; MARCEL, Sébastien; ANJOS, André. *COHFACE*. 2016. Dostupné z DOI: 10.34777/ff3f-ba56.
26. LATHI, B P. *Linear Systems and Signals*. 2. vyd. New York, NY: Oxford University Press, 2004. The Oxford Series in Electrical and Computer Engineering.
27. JAIN, Vidit; LEARNED-MILLER, Erik. *Fddb: A Benchmark for Face Detection in Unconstrained Settings*. 2010. Tech. zpr., UM-CS-2010-009. University of Massachusetts, Amherst.
28. *Fddb: Face Detection Data Set and Benchmark* [online]. [cit. 2023-04-23]. Dostupné z: <http://vis-www.cs.umass.edu/fddb/samples/>.
29. BOCCIGNONE, Giuseppe; CONTE, Donatello; D'AMELIO, Alessandro; CUCULO, Vittorio; GROSSI, Giuliano; LANZAROTTI, Raffaella. An Open Framework for Remote-PPG Methods and their Assessment. *IEEE Access*. 2020, s. 1–1. Dostupné z DOI: 10.1109/access.2020.3040936.
30. HEUSCH, Guillaume; ANJOS, André; MARCEL, Sébastien. *A Reproducible Study on Remote Heart Rate Measurement*. 2017. Dostupné z arXiv: 1709.00962 [cs.CV].
31. BOBBIA, S.; MACWAN, R.; BENEZETH, Y.; MANSOURI, A.; DUBOIS, J. Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognition Letters*. 2017.

32. GOPRO. *Hero4 Silver: User Manual* [online]. [B.r.]. [cit. 2023-04-17]. Dostupné z: [https://gopro.com/content/dam/help/hero4-silver/manuals/UM\\_H4Silver\\_ENG\\_REVA\\_WEB.pdf](https://gopro.com/content/dam/help/hero4-silver/manuals/UM_H4Silver_ENG_REVA_WEB.pdf).
33. *OpenCV Cascade Classifier* [online]. [cit. 2023-04-23]. Dostupné z: [https://docs.opencv.org/3.4/db/d28/tutorial\\_cascade\\_classifier.html](https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html).
34. DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. 2005, sv. 1, 886–893 vol. 1. Dostupné z DOI: 10.1109/CVPR.2005.177.
35. KING, Davis E. *Max-Margin Object Detection*. 2015. Dostupné z arXiv: 1502.00046 [cs.CV].
36. BAZAREVSKY, Valentin; KARTYNNIK, Yury; VAKUNOV, Andrey; RAVEENDRAN, Karthik; GRUNDMANN, Matthias. *BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs*. 2019. Dostupné z arXiv: 1907.05047 [cs.CV].
37. LIU, Wei; ANGUELOV, Dragomir; ERHAN, Dumitru; SZEGEDY, Christian; REED, Scott; FU, Cheng-Yang; BERG, Alexander C. SSD: Single Shot MultiBox Detector. In: *Computer Vision – ECCV 2016*. Springer International Publishing, 2016, s. 21–37. Dostupné z DOI: 10.1007/978-3-319-46448-0\_2.
38. KING, Davis E. Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research*. 2009, roč. 10, s. 1755–1758.
39. *Dlib CNN* [online]. [cit. 2023-04-23]. Dostupné z: <http://blog.dlib.net/2016/10/easily-create-high-quality-object.html>.
40. ZHANG, Kaipeng; ZHANG, Zhanpeng; LI, Zhifeng; QIAO, Yu. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. *IEEE Signal Processing Letters*. 2016, roč. 23, č. 10, s. 1499–1503. Dostupné z DOI: 10.1109/LSP.2016.2603342.
41. DENG, Jiankang; GUO, Jia; ZHOU, Yuxiang; YU, Jinke; KOTSIA, Irene; ZAFEIRIOU, Stefanos. *RetinaFace: Single-stage Dense Face Localisation in the Wild*. 2019. Dostupné z arXiv: 1905.00641 [cs.CV].
42. DAHMANI, Djamila; CHEREF, Mehdi; LARABI, Slimane. Zero-sum game theory model for segmenting skin regions. *Image and Vision Computing*. 2020, roč. 99, s. 103925. ISSN 0262-8856. Dostupné z DOI: <https://doi.org/10.1016/j.imavis.2020.103925>.
43. ABDULLAH-AL-WADUD, M.; SHOYAIB, Mohammad; CHAE, Oksam. A Skin Detection Approach Based on Color Distance Map. *EURASIP Journal on Advances in Signal Processing*. 2009, roč. 2008, č. 1. Dostupné z DOI: 10.1155/2008/814283.

44. SMITH, Alvy Ray. Tint fill. In: *Proceedings of the 6th annual conference on Computer graphics and interactive techniques - SIGGRAPH '79*. ACM Press, 1979. Dostupné z DOI: 10.1145/800249.807456.
45. COMANICIU, D.; MEER, P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002, roč. 24, č. 5, s. 603–619. Dostupné z DOI: 10.1109/34.1000236.
46. ROTHER, Carsten; KOLMOGOROV, Vladimir; BLAKE, Andrew. "Grab-Cut". In: *ACM SIGGRAPH 2004 Papers*. ACM, 2004. Dostupné z DOI: 10.1145/1186562.1015720.
47. KARTYNNIK, Yury; ABLAVATSKI, Artsiom; GRISHCHENKO, Ivan; GRUNDMANN, Matthias. *Real-time Facial Surface Geometry from Monocular Video on Mobile GPUs*. 2019. Dostupné z arXiv: 1907.06724 [cs.CV].
48. XU, Han; SARKAR, Abhijit; ABBOTT, A. Lynn. *Color Invariant Skin Segmentation*. arXiv, 2022. Dostupné z DOI: 10.48550/ARXIV.2204.09882.
49. TADELIS, Steven. *Game Theory: An Introduction*. Princeton University Press, 2013. ISBN 978-0691129082.
50. MÜLLER, Andreas C.; GUIDO, Sarah. *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, Inc., 2016. ISBN 9781449369415.
51. HAAN, Gerard de; JEANNE, Vincent. Robust Pulse Rate From Chrominance-Based rPPG. *IEEE Transactions on Biomedical Engineering*. 2013, roč. 60, č. 10, s. 2878–2886. Dostupné z DOI: 10.1109/TBME.2013.2266196.
52. PILZ, Christian; ZAUNSEDER, Sebastian; KRAJEWSKI, Jarek; BLAZEK, Vladimir. Local Group Invariance for Heart Rate Estimation from Face Videos in the Wild. In: 2018. Dostupné z DOI: 10.1109/CVPRW.2018.00172.
53. WANG, Wenjin; STUIJK, Sander; HAAN, Gerard de. A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation. *IEEE Transactions on Biomedical Engineering*. 2016, roč. 63, č. 9, s. 1974–1984. Dostupné z DOI: 10.1109/TBME.2015.2508602.
54. HAAN, G de; LEEST, A van. Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiological Measurement*. 2014, roč. 35, č. 9, s. 1913–1926. Dostupné z DOI: 10.1088/0967-3334/35/9/1913.





## Obsah přiloženého média

	readme.txt.....	stručný popis obsahu média
	src	
	impl.....	zdrojové kódy implementace
	thesis.....	zdrojová forma práce ve formátu $\text{\LaTeX}$
	text	
	thesis.pdf.....	text práce ve formátu PDF