

# Posudek práce

předložené na Ústavu Aplikované Informatiky Přírodovědecké fakulty JU

## posudek oponenta bakalářské práce

**Autor:** František Dolejš

**Název práce:** Forenzní analýza grafických souborů a video souborů  
**Studijní obor: zadejte obor** Aplikovaná informatika  
**Datum odevzdání:** 4. 12. 2014

**Jméno a tituly oponenta:** kpt. Ing. Radovan Hlaváček

**Pracoviště:** Policie ČR, Krajské ředitelství policie Královéhradeckého kraje, Odbor Kriminologické Techniky a Expertiz, Oddělení Počítačových expertiz, expert v odvětví kriminalistika, analýza dat a zkoumání nosičů dat, metrolog

**Kontaktní e-mail:** radovan.hlavacek@pcr.cz

**Odborná úroveň práce:** výborná

**Věcné chyby:** téměř žádné

**Výsledky:** velmi dobré

**Rozsah práce:** standartní

**Grafická, jazyková a formální úroveň:** velmi dobrá

**Tiskové chyby:** téměř žádné

**Celková úroveň práce:** velmi dobrá

### **Slovní vyjádření, komentáře a připomínky vedoucího práce:**

Zvolené téma bakalářská práce je pro praktické využití ve forenzní praxi velice zajímavé a v současné době žádané. Kriminalita v oblasti pornografie a obzvláště dětské pornografie rok od roku stoupá. Stoupá i objem dat a náročnost zkoumání datových nosičů potenciálně podezřelých z výskytu pornografie. Často jsou využívány pro ukládání závadných dat jak pevná (lokální) datová úložiště, tak i vzdálená úložiště (online služby). Manuální prohlížení obsahu souborů je velice náročné a zdouhavé. V případě vzdálených úložišť s velkým počtem uživatelů (veřejná úložiště velkých subjektů) je manuální prohlížení obsahu souborů prakticky nerealizovatelné. Obzvláště pro velké objemy zpracovávaných dat je vhodné a výhodné použít automatických metod prohledávání. Vytváření vlastních novátorských metod k vyhledávání a hodnocení obsahu souborů by bylo velice prospěšné, ale v rámci bakalářské práce jednoho člověka nadmíru náročné. Využití stávajících a funkčních vyhledávacích a hodnotících algoritmů, které používají velké subjekty vlastníci veřejná úložiště (Microsoft, Google, Facebook, atd..) hodnotím jako vhodné. V současné době je na trhu softwarových produktů velice málo praktických řešení této problematiky. Pokud se vyskytují, mají výrazně omezené možnosti detekce pornografie.

Autor ve své práci uvedl správně základní právní skutečnosti, v současné době kodifikující problematiku pornografie i dětskou pornografii. Dále základní zásady zajišťování dat pro forenzní účely, což je klíčové. Bez správně zajištěných dat je následné zkoumání bezcenné. Popsal obecné i praktické přístupy velkých společností poskytující online služby, jako Microsoft, Google, Facebook, Twitter. Metody počítačového vidění zde popsané mají praktické využití i v jiných oblastech, např. strojním a elektrotechnickém průmyslu. V práci jsou velice podrobně popsány metody rozpoznávání lidské kůže na základě hodnocení histogramu barevného modelu kůže YCbCr a detekce objektů Viola-Jones, metoda EmguCV, včetně základních klasifikačních algoritmů. Dále jsou zde popsány i jiné metody detekce lidské kůže a možná vylepšení. Jsem přesvědčen, že podrobný popis zde uvedených metod přesahuje odbornou stránku bakalářské práce a hodnotím velice kladně. V práci jsou popsány metody rozpoznání postavy dítěte na obrázku metodou poměrovou, porovnání částí těla a rysů v obličeji, což je důležité pro základní rozpoznání dětské pornografie.

Autor navrhnul jednoduchou aplikaci „Detektor“ pro vyhodnocování závadnosti obrázků s využitím metody detekce obličeje EmguCV a detekce barvy kůže, tak jak popisuje teorie. Výsledný PDF s výsledkem rozpoznávání, obsahuje seznam obrázků, případně načtené EXIF informace z testovaných obrázků a velice důležitou sumu souborů SHA1, jež jednoznačně identifikuje neměnnost obsahu souboru. Autor provedl základní statistické vyhodnocení výsledků, které aplikace produkuje, což je velice důležité pro další práci na aplikaci. Předpokládám, že to není konečný produkt, ale velice dobrý základ pro další práci. V rámci tohoto projektu hodnotím úspěšnost funkce aplikace jako velice dobrou. Jako částečný nedostatek hodnotím nutnost výskytu obrázků v jednom adresáři. Navrhuji doplnit rekurzivní načítání obrázků v celé adresářové struktuře. Dále také doplnit aplikaci o identifikaci obrázkových souborů na základě hlavičky souborů aby bylo možné detekovat obrázkové soubory se změněnou příponou souboru. Navrhuji rovněž rozšířit okruh zkoumaných typů souborů. Z hlediska datového výstupu aplikace bych navrhoval doplnit o export výpisu do dalších typů souborů, jako např. HTML, CLS, XLSX. Vycházím z vlastní praxe a požadavků dožadujících orgánů v případě zkoumání pornografie. Autorem navrhovaný detektor bradavek by jistě zvýšil schopnosti detekce aplikace a snížil její chybovost. Dále navrhuji, zda by bylo možné do aplikace implementovat funkci nastavení rozsahu citlivosti detekce rozpoznávání lidské kůže případně detekce poměrů rysů v obličeji, vs. počet typů detekovaných souborů. Čím užší bude rozsah detekce, tím méně typů aplikace vyprodukuje, zato



s větší přesností a naopak. Obdobná funkce se vyskytuje v programech detekujících papírní linie otisků prstů.

Závěrem bych rád upozornil na drobný nedostatek grafické podoby bakalářské práce, kdy odkazy na obrázky nejsou na stejné stránce jako samotné obrázky. Je to hůře čitelné. Doporučil bych také „zlidštit“ popis aplikace „Detektor“, je velice technicky podrobný. Veškeré výhrady, které k práci mám, jsou jen upozorněním na drobné nedostatky a doporučuji je brát jako námět k dalšímu rozvoji aplikace. Tato práce, v mnohém převyšuje svou technickou úroveň, bakalářskou práci.

**Slovní hodnocení:**

**Uvedená práce splňuje požadavky na bakalářskou práci a doporučuji ji k obhajobě, s celkovým hodnocením „velmi dobře“.**

**Případné otázky při obhajobě a náměty do diskuze:**

viz. výše

**Otázky**

Žádné další otázky

**Práci doporučuji uznat jako bakalářskou.**

**Navrhuji hodnocení stupněm: velmi dobře**

V Hradci Králové dne 22. 12. 2014

kpt. Ing. Radovan Hlaváček  
oponent bakalářské práce

