**Rozšírené zadanie diplomovej práce**

**Názov:**Rozpoznávanie a klasifikácia fontov písma s využitím vlastného expertného systému

**Fakulta/Univerzita:**Prírodovedecká fakulta UPJŠ v Košiciach

**Stredisko:**ÚINF –Ústav informatiky

**Akademický rok:**2020/2021

**Vedúci:**RNDr. Ľubomír Antoni PhD.

**Konzultant:**doc. RNDr. Stanislav Krajči, PhD.

**Autor:**Bc. Dominik Dorčák

**Ciele práce:**

1. Analyzovať reprezentáciu a spôsoby uchovávania fontov v počítačových systémoch
2. Navrhnúť množinu vlastností pre možnosť odlišovania rôznych fontov písma.
3. Implementovať a otestovať expertný systém na klasifikáciu fontov písma pomocou odpovedí užívateľa.

**Literatúra:**

1. Kuzma, M., Andrejková, G.: Predicting user’s preferences using neural networks and psychology models. [Applied Intelligence](https://link.springer.com/journal/10489), 2016; 44(3): 526-538.
2. Zramdimi, A, Ingold, R.: Optical font recognition using typographical features, 1998; 20(8): 877-882.
3. Janč, R.: Rozpoznávanie typu fontu tlačeného písma [ Diplomová práca ] Univerzita Komenského, Fakulta matematiky, fyziky a informatiky, Katedra informatiky. 1998.

**Rozšírená anotácia:**

Font je skupina písmen, číslic a interpunkčných znamienok rovnakého grafického prevedenia, štýlu a hrúbky. Existujú rôzne typy fontov. Môžeme ich deliť do viacerých skupín na základe rozličných vlastností. Podľa vzhľadu ich delíme na pätkové, bezpätkové a typu písací stroj. Podľa šírky obálky písmena ich delíme na proporcionálne a neproporcionálne. Z hľadiska reprezentácie v počítači ich delíme na bitmapové a vektorové [3].

V súčasnosti je bežné spracovanie dokumentov pomocou počítačov. Pri spracovaní dokumentov je mnohokrát potrebné vystačiť si s materiálmi v tlačenej papierovej podobe. V prípade, ak požadujeme vytvoriť diplom, vizitku, či zmluvu s identickým typom písma, aký je použitý na vytlačenom dokumente, znalosť typu písmu na dokumente je pre nás dôležitou informáciou.

Zaujímavým prípadom reprezentácie fontov je takzvaný metafont používaný v typografických systémoch na báze systému TeX. Z tejto reprezentácie je možné taktiež odčítať niekoľko vlastností fontu.

Pre klasifikáciu fontov písma, sa v literatúre stretávame s riešeniami vo forme OCR [3] alebo implementácie klasifikačnej úlohy s využitím neurónových sietí, či genetických algoritmov [1]. Úspešnosť takýchto riešení nie je 100 %, čo nás motivuje k nájdeniu vlastností fontov, ktoré nám dané typy písma čo najpresnejšie rozlíšia.

Väčšina existujúcich riešení berie ako vstup obrázok z ukážkou textu. Myšlienkou tejto práce je detegovať font na základe vlastností odpozorovaných používateľom. Našou úlohou je analyzovať reprezentáciu a spôsoby uchovávania fontov v počítačových systémoch, navrhnúť množinu vlastností pre možnosť odlišovania rôznych fontov písma a implementovať deterministický expertný systém na klasifikáciu fontov písma pomocou odpovedí užívateľa na zadané otázky. Expertný systém otestujeme na rôznych typoch dokumentov s využitím odpovedí užívateľov.