

Zagreb, 9. ožujka 2017.

ZAVRŠNI ZADATAK br. 5187

Pristupnik: **Kristijan Vulinović (0036483282)**

Studij: Računarstvo

Modul: Računarska znanost

Zadatak: Ekstrakcija tablica na skeniranim dokumentima

Opis zadatka:

Problem očitavanja sadržaja skeniranih dokumenata javlja se u mnogim područjima ljudskoga rada. Automatizacija tog postupka može značajno smanjiti vrijeme koje se troši na ručno očitavanje. Problem se pri tome može razložiti na niz potproblema, ovisno o strukturi informacija koje su prisutne u dokumentu.

U okviru završnog rada potrebno je istražiti i opisati postupke precizne detekcije tablica. Zadaća je kvalitetno detektirati položaj tablice te izolirati svaku od ćelija tablice, kako bi se njihov sadržaj potom mogao obraditi nekim drugim postupkom. Pri tome treba voditi računa da se u okolini tablice mogu zateći različite smetnje te da ćelije tablice mogu biti popunjene različitim sadržajima. Potrebno je pripremiti odgovarajući skup uzoraka za učenje, napisati programsku implementaciju odabranog postupka u programskom jeziku Java, provesti vrednovanje implementiranog postupka te komentirati dobivene rezultate.

Radu je potrebno priložiti algoritme, izvorne kodove i rezultate uz potrebna objašnjenja i dokumentaciju. Citirati korištenu literaturu i navesti dobivenu pomoć.

Zadatak uručen pristupniku: 10. ožujka 2017.

Rok za predaju rada: 9. lipnja 2017.

Mentor:



Doc. dr. sc. Marko Čupić

Djelovođa:



Doc. dr. sc. Tomislav Hrkać

Predsjednik odbora za
završni rad modula:



Prof. dr. sc. Siniša Srbljić