

Práca skúma problém detegovania anonymizovaných častí v PDF dokumentoch. Preskúmané boli rôzne prístupy na detekciu, primárne analýza obrazu a s ňou spojené rôzne algoritmy počítačového videnia. Najlepší z týchto prístupov sme implementovali a vyhodnotili na testovacích dátach. Výsledky ukázali, že implementovaný prístup dosiahol vysokú presnosť a predbehol iné prístupy aj vzhľadom na efektivitu. Tento výskum prispieva k rozvoju nástrojov pomáhajúcich analyzovať dokumenty, ktoré môžu byť aplikované v rôznych právnych či finančných oblastiach na zaručenie ochrany dát v súlade s reguláciami.