

Kamerové systémy dnes slouží různým účelům, od bezpečnosti k monitorování dopravy a marketingu. Nicméně, s rostoucím množstvím kamer se stává manuální sledování videa příliš pracné. V posledních letech se hodně vývoje umělé inteligence zaměřilo na automatické zpracování videa a následný výstup požadovaných upozornění a statistik. Tato práce zkoumá nejmodernější modely hlubokého učení pro detekci objektů v bezpečnostním videu a podrobně se zabývá SSD architekturou. Naším hlavním cílem je zvýšit výkon SSD architektury aktualizací vnitřní sítě extrahující tzv. feature mapy. V práci jsou navrženy možnosti nahrazení původního VGG modelu pomocí nejnovějších klasifikačních sítí ResNet, Xception a NASNet. Experimentálně jsme zjistili, že model ResNet50 nabízí nejlepší kompromis mezi rychlostí a přesností. Tento model zároveň výrazně překonává VGG. Po zavedení řady modifikací do sítě Xception se nám povedlo dorovnat výkon ResNetu. Kromě vylepšení architektury také analyzujeme vztah mezi SSD a množstvím detekovaných tříd a jejich výběrem. Také jsme navrhli a implementovali nový detektor, který využívá temporální kontext snímku pro detekci objektů. Tento detektor pracuje v reálném čase a se zvýšenou přesností.