

Automatická detekce objektů

Prostor je sledovaný několika kamerami, jejichž zorná pole se překrývají. To umožňuje lepší pokrytí sledovaného prostoru než s použitím jedné kamery, ale vyžaduje to složitější analýzu. Proces získání trajektorií objektů lze rozdělit na několik kroků:

1. Detekce objektů ve videích s pomocí neuronové sítě.
2. Převod pozorovaných pozic do globální souřadné soustavy.
3. Identifikace pozorovaných pozic s objekty a sledování objektů v čase.
4. Analýza výsledných trajektorií

Tyto kroky jsou detailněji popsány v následujících odstavcích.

1. Detekce objektů

Objekty jsou ve videích detekovány s pomocí neuronové sítě. Použili jsme síť detekující boxy architektury RetinaNet, ale lze použít jakoukoliv síť ze které se získat pozice sledovaných objektů.

2. Převod do globální souřadné soustavy

Pro syntézu dat z více kamer je třeba převést souřadnice objektů z obrazového prostoru (dvě souřadnice v pixelech) do globálního prostoru stejného pro všechny kamery (tři souřadnice v metrech). To lze provést pouze se znalostí optických vlastností, poloh a orientací kamer. Parametry popisující tyto vlastnosti jsou souhrnně označovány jako kalibrace. Kalibrace musí být předem vypočtena pro použitou soustavu kamer.

S pomocí kalibrace je možné převést obrazové souřadnice objektů na paprsky v prostoru které vedou od kamery směrem k objektu. Vzdálenost objektu z jedné kamery určit nelze. Pokud však předpokládáme že se všechny objekty pohybují v rovině země ($Z=0$), tak pozici $[x, y, 0]$ určit lze.

Kromě polohy objektů ve světových souřadnicích jsou vypočteny také nejistoty reprezentované jako matice kovariancí. Předpokládáme, že detekce objektu má konstantní neurčitost v pixelech, a tuto neurčitost převedeme do globálního prostoru. To nám umožňuje v dalším kroku pracovat efektivně s objekty blízko u kamery i daleko od kamery, kde se nejistota v určení polohy může mnohonásobně lišit.

3. Identifikace pozorování s objekty

Vzhled objektů nebereme v potaz, pouze jejich polohy. Je tedy třeba odhadnout, která pozorování jsou generována kterými objekty. To je prováděno Bayesovským schématem na základě pozic a jejich kovariancí. Pokud je likelihood toho že dané pozorování bylo generováno daným objektem větší než práh, je pozorování k objektu asociováno, a je použito pro zpřesnění jeho polohy. Polohy objektů v čase jsou propagovány v čase s pomocí Kalmanova filtru. Předpokládáme, že objekty se pohybují s konstantní rychlostí, a objekty jsou inicializovány s velkou nejistotou v rychlosti. Kalmanův filtr umožňuje sledovat i dočasně skryté objekty, což je běžné například pro dvě osoby které jdou blízko u sebe.

Protože jsou asociace pozorování a objektů prováděny čistě geometricky, kamery musí být dobře kalibrované, a jejich časová synchronizace musí být přesně určena. Jinak dochází ke zdvojování objektů, protože nedochází ke korektní asociaci.

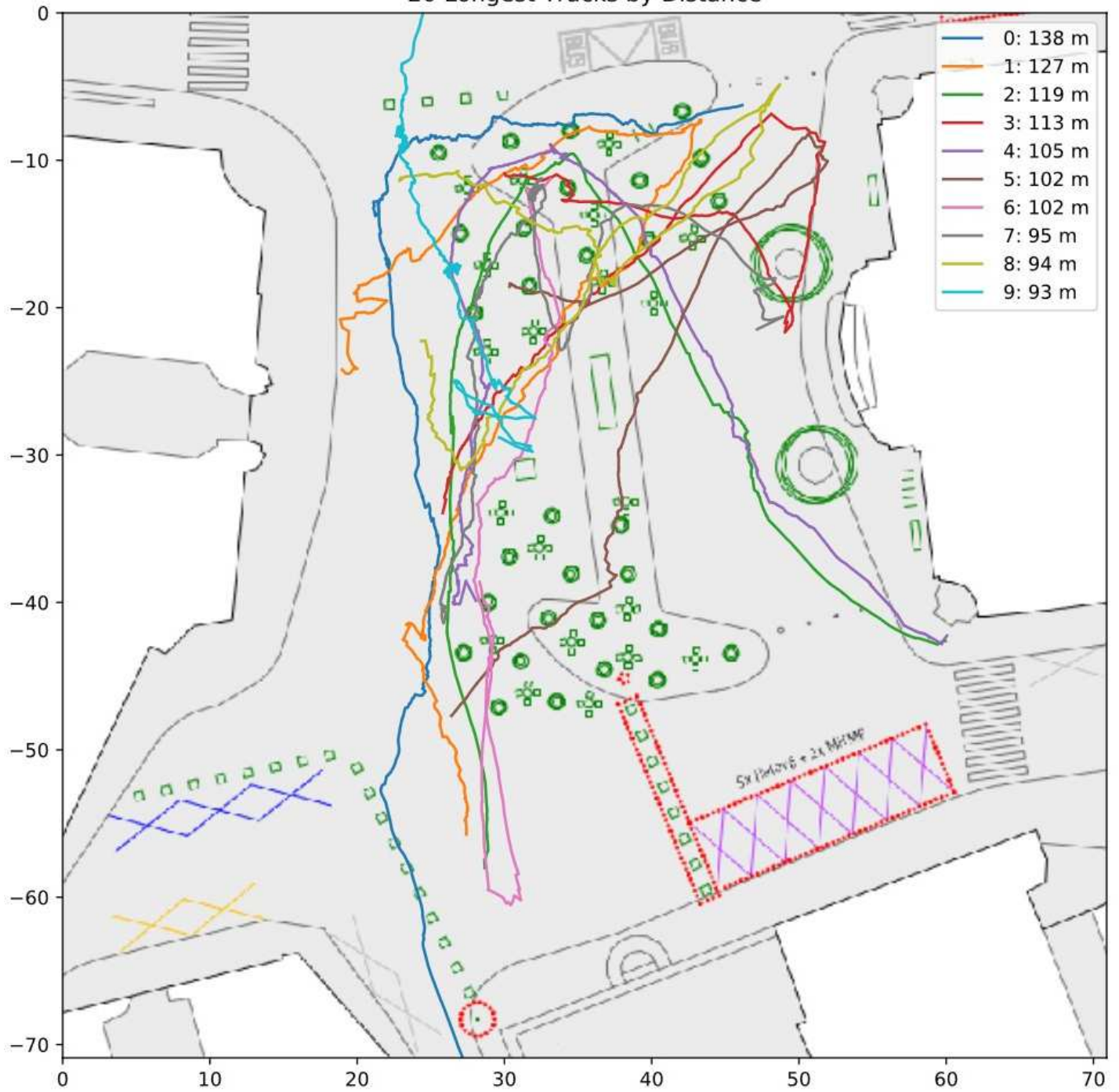
4. Analýza výsledných trajektorií

Výsledkem předchozích bodů jsou trajektorie objektů ve světových souřadnicích. Tyto trajektorie lze dále zpracovávat. Lze provádět například:

- Filtraci podle délky trajektorie
- Filtraci podle doby, kterou osoba strávila na sledovaném místě
- Filtraci podle zón, ve kterých se osoba zdržovala
- Výpočet, kolik osob je průměrně v jaké části sledovaného místa

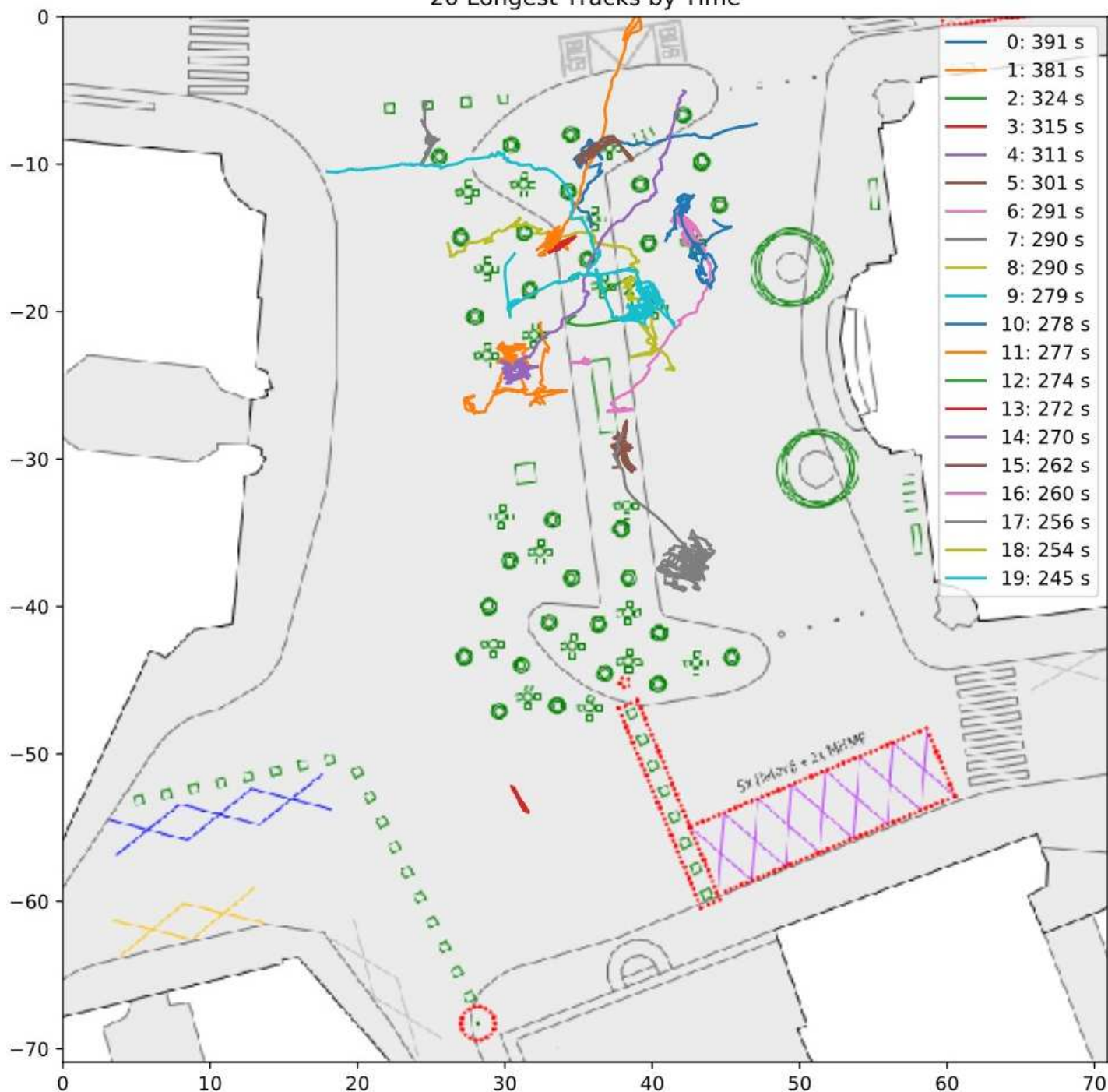
Výsledky mohou být jednoduše vizualizovány na mapě; stačí znát převod mapových souřadnic do globální souřadné soustavy. Například, vizualizace 10 nejdelších sledovaných trajektorií z Mariánského náměstí:

20 Longest Tracks by Distance



Visualizace 10 trajektorií, které jsou časově nejdelší:

20 Longest Tracks by Time



Přesnost detekce objektů lze odhadnout na základě předpokladu, že osoby odcházejí a přicházejí pouze do/ze známých zón. Pokud se objekt objeví uprostřed sledované oblasti, došlo zřejmě k "roztržení" trajektorie. Tyto chyby mohou být zřejmě pouze minimalizovány, a je třeba s jejich výskytem počítat při analýze výsledků.