

Robusno vizualno prepoznavanje mesta uporabom dubokih reprezentacija i uparivanja slijedova slika



Jurica Maltar, mag. math.

(jurica.maltar@fer.hr)

mentor: izv. prof. dr. sc. Ivan Marković, izv. prof. dr. sc. Domagoj Matijević
Sveučilište u Zagrebu Fakultet elektrotehnike i računarstva

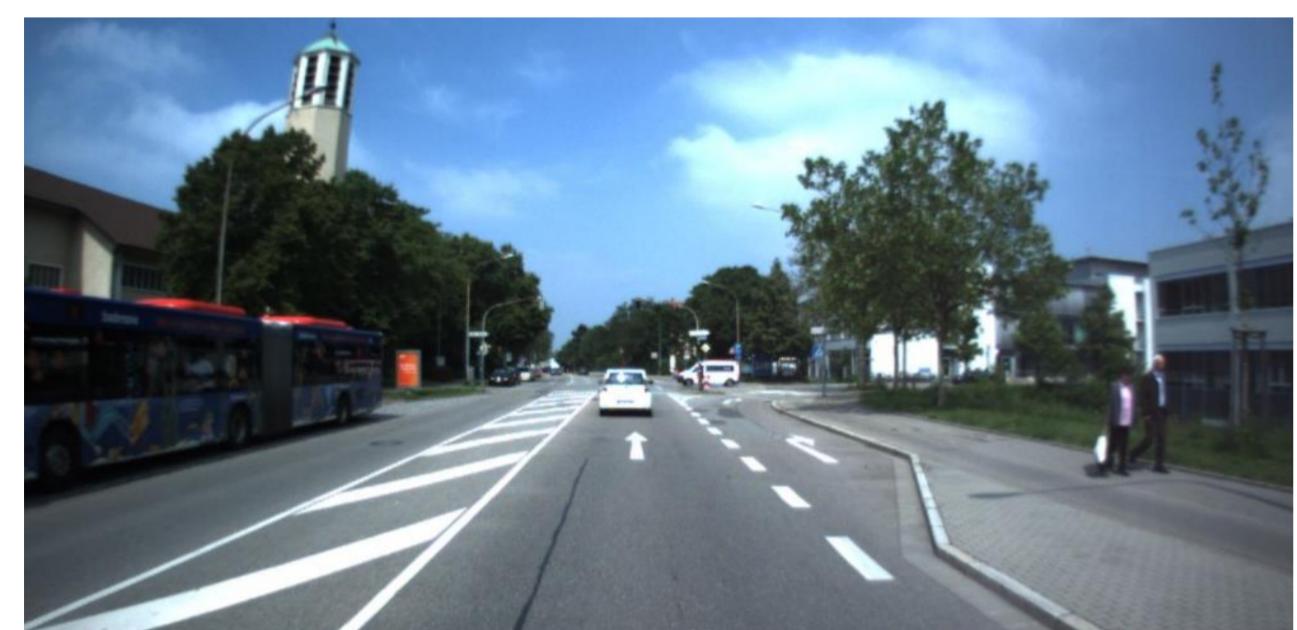


1. Uvod

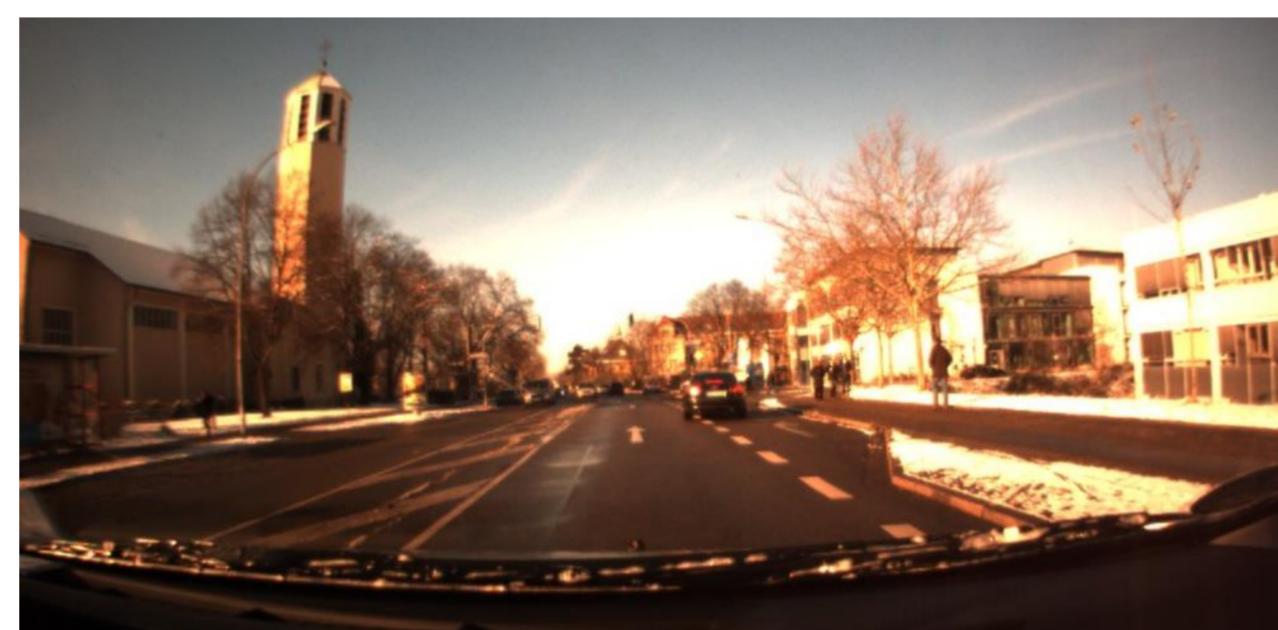
Vozilo koje izvodi *istovremenu lokalizaciju i kartiranje* (abbr. **SLAM**) gibajući se neizbjegivo akumulira pogrešku. Međutim, ako se vozilo nađe na već otprije viđenom mjestu, moguće je *zatvoriti petlju* te na taj način umanjiti pogrešku. Zatvaranje petlje istovjetno je prepoznavanju otprije viđenog mesta te se takav problem naziva **prepoznavanje mesta**. Ako je senzorski modalitet vizualni senzor, govorimo o *vizualnom prepoznavanju mesta*.

2. Opis problema

Upravo snimljenu sliku – **sliku upita** I_{Q_i} – uspoređujemo s otprije snimljenim slikama – **referentnim slikama** $I_{D_j}, \forall j$ – te pronalazimo najprikladniju, odn. onu sliku koja prikazuje isto mjesto. Dvije različite slike istog mesta mogu se razlikovati i zbog različitih gledišta kamere i zbog različitih okolišnih uvjeta. Komponente takvog sustava su: **metoda uparivanja mesta i vizualna reprezentacija mesta**.



I_{Q_i}

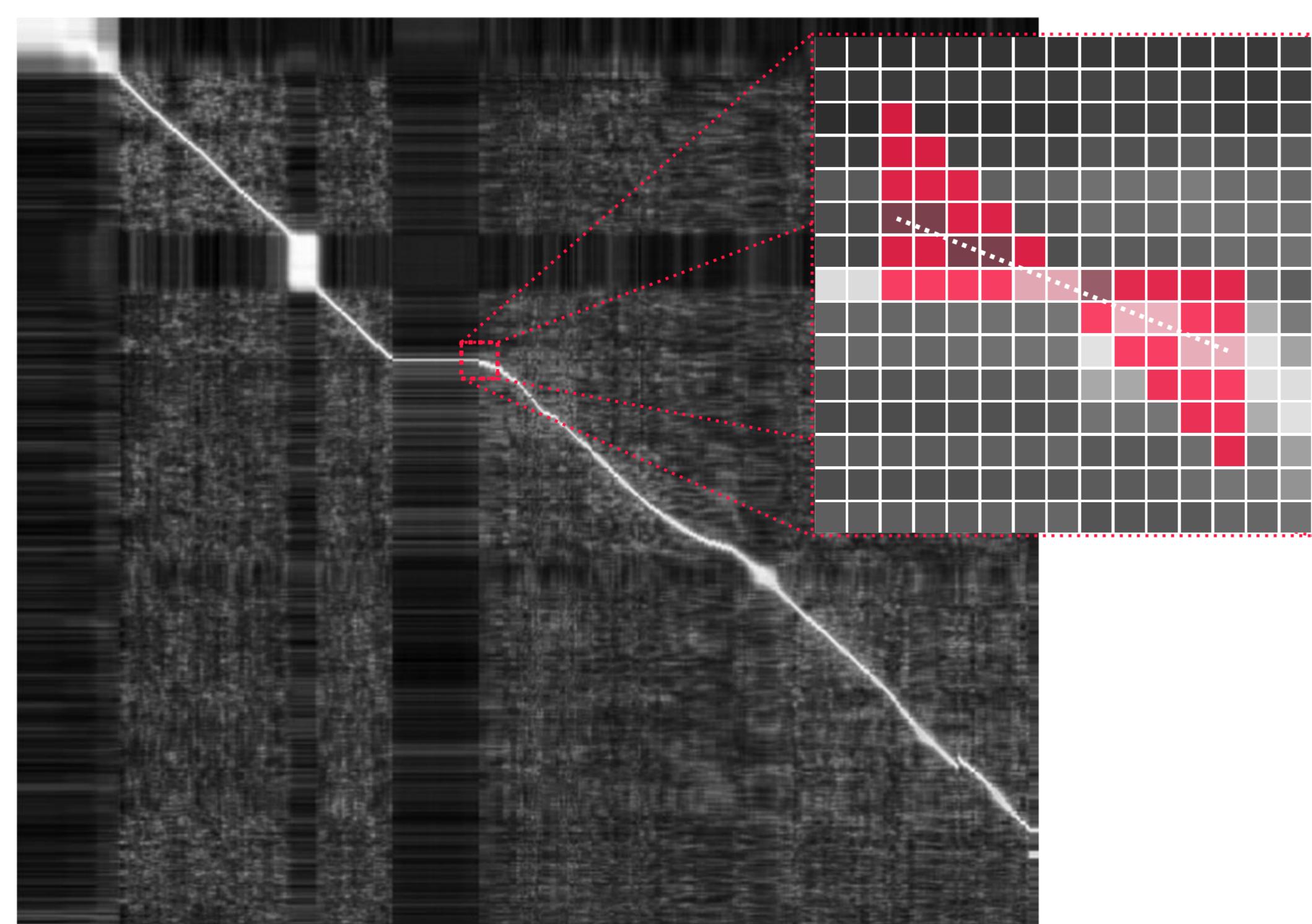


I_{D_j}

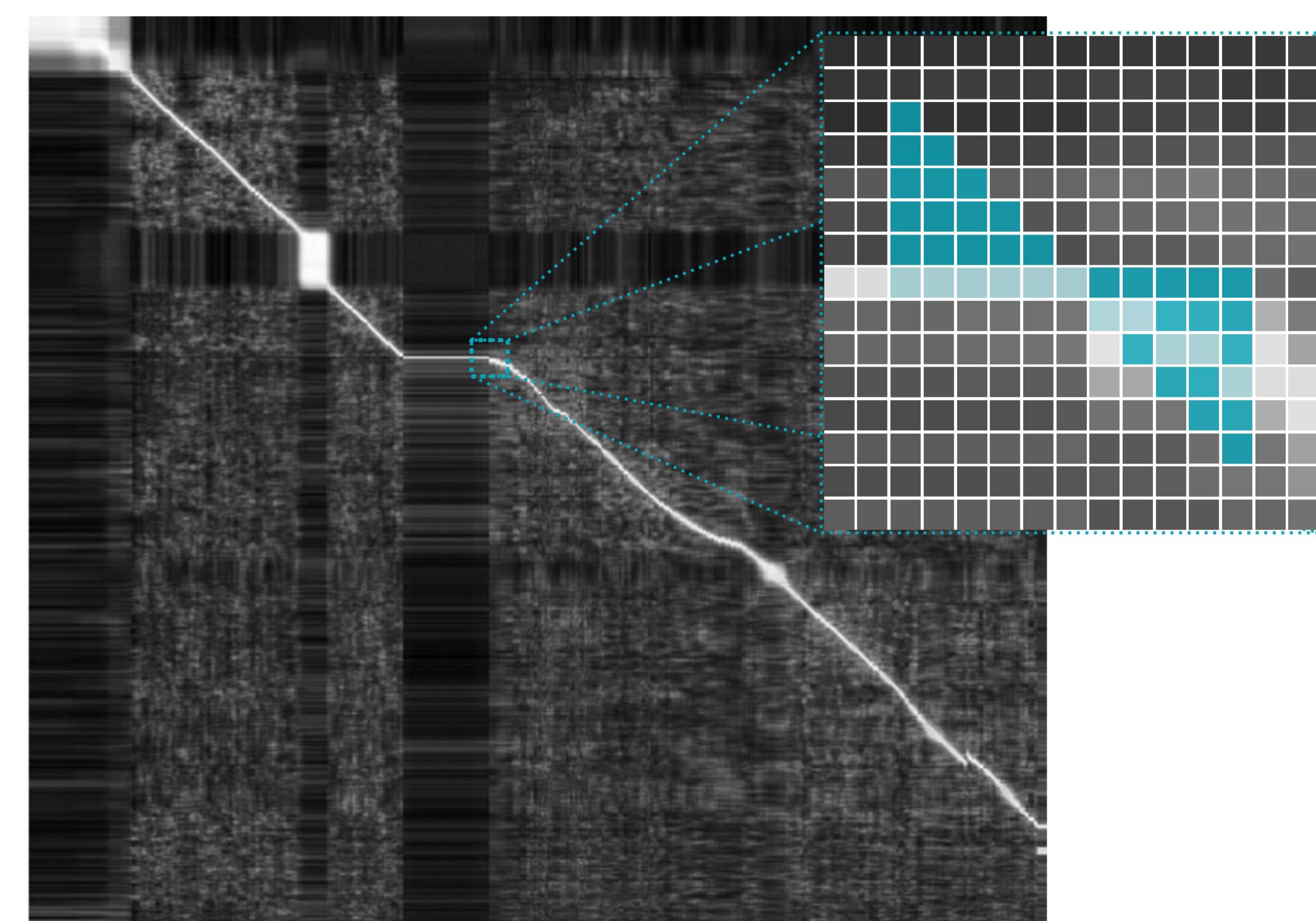
3. Metodologija

U novije se vrijeme kao *globalni opisnik slike* koriste *mapa značajki* izlučene *dubokim konvolucijskim neuronskim mrežama* (abbr. **DCNN**). Korištenjem softmax regresije fino podešavamo arhitekturu neuronske mreže. Dodatno, *odabirom značajki zasnovanim na zajedničkoj informaciji* odabiremo bolje te izuzimamo lošije značajke.

Globalni opisnik z_{Q_i} slike upita I_{Q_i} naivno možemo usporediti s opisnicima referentnih slika $z_{D_j}, \forall j$ pa odabrati najsličniju sliku. Napredniji je pristup koristiti **slijed** slika nastalih neposredno prije (i nakon ukoliko se radi o tzv. *offline* metodi uparivanja) snimanja trenutne slike. Taj se slijed naziva i *lokalno susjedstvo*. **SeqSLAM** računa prikladnost uparivanja slika I_{Q_i} i I_{D_j} uvezši u obzir isključivo linearnu korelaciju u brzini vozila pri različitim obilascima istog mesta.



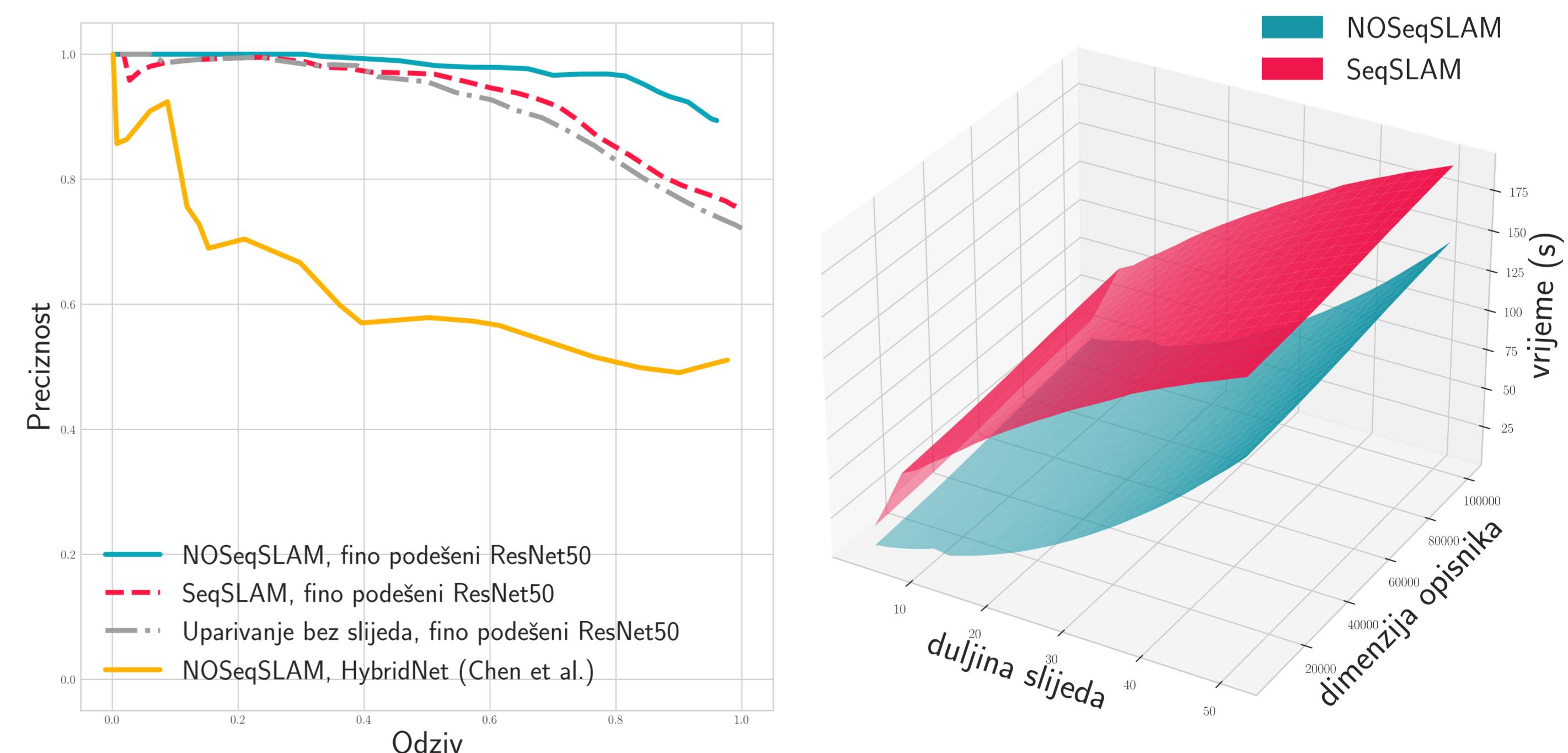
U stvarnosti se vozilo giba različitim brzinama pri obilasku istog mesta. Naš pristup – **NOSeqSLAM** – uzima u obzir takve scenarije te je poopćenje prethodnog pristupa. Nelinearna se korelacija u brzinama obuhvaća izvršavanjem algoritma *najkraćeg puta iz jednog izvora* nad strukturu grafom.



4. Rezultati

Najčešća je kvantitativna ocjena evaluacije metode uparivanja mesta površina pod *odziv-preciznost* krivuljom (donja lijeva slika). Za skup podataka "The Oxford RobotCar Dataset" vidimo kako NOSeqSLAM postiže bolje kvantitativne rezultate od suparničke metode SeqSLAM, a obje metode postižu bolje rezultate od naivnog uparivanja bez slijeda. Također je vidljivo kako fino podešeni **ResNet** nadmašuje suparničku vizualnu reprezentaciju **HybridNet**. Rezultati dobiveni nepodešenom mrežom znatno su lošiji pa ih izostavljamo.

Iako je *asimptotsko vrijeme izvršavanja* nešto bolje za SeqSLAM, za praktične duljine slijeda i praktične dimenzije opisnika, empirijsko je vrijeme izvršavanja NOSeqSLAM-a kraće (donja desna slika).



5. Zaključak

Pri vizualnom prepoznavanju mesta suočavamo se s promjenama vizualne reprezentacije nastale zbog različitih gledišta kamere i okolišnih uvjeta. Uparivanjem slijedova slika postižemo bolje kvantitativne rezultate gdje NOSeqSLAM, kao poopćenje, nadmašuje SeqSLAM. Također, naša fino podešena vizualna reprezentacija uvelike pospješuje rezultate.