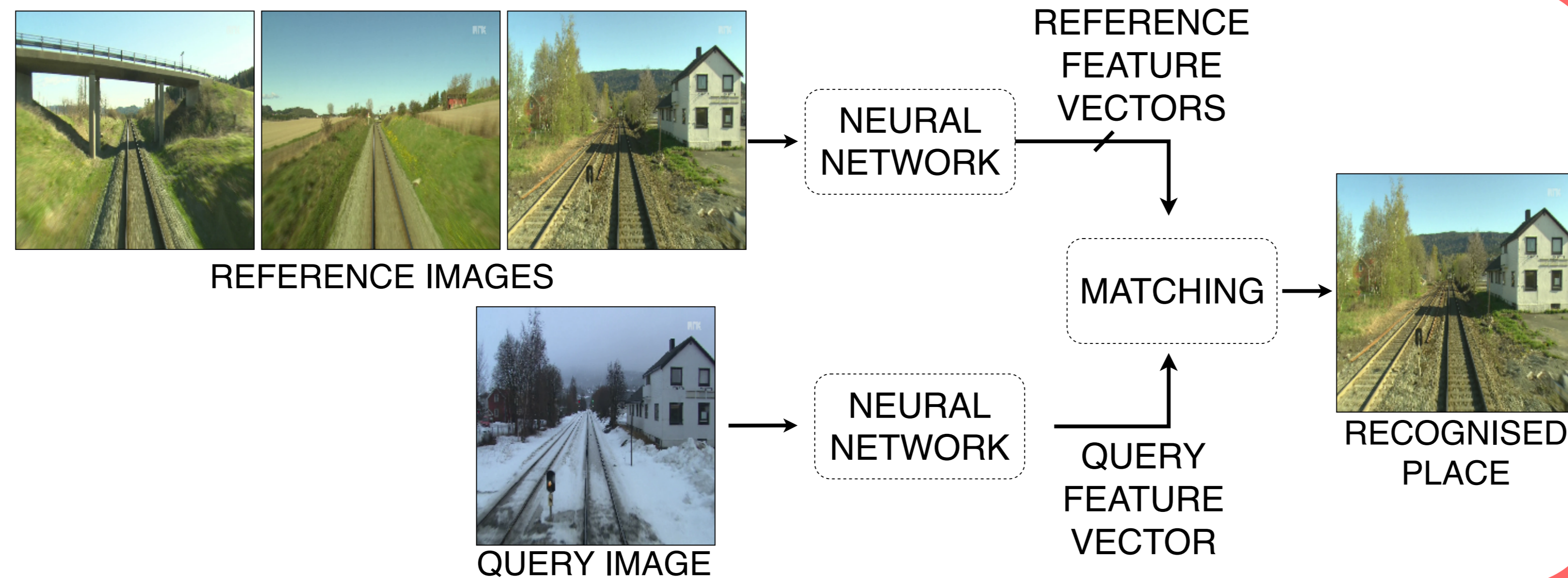


Daniel Olid, José M. Fácil and Javier Civera
 The authors are with the I3A, Universidad de Zaragoza, Spain
 danielolid94@yahoo.es, jmfácil@unizar.es, jcivera@unizar.es

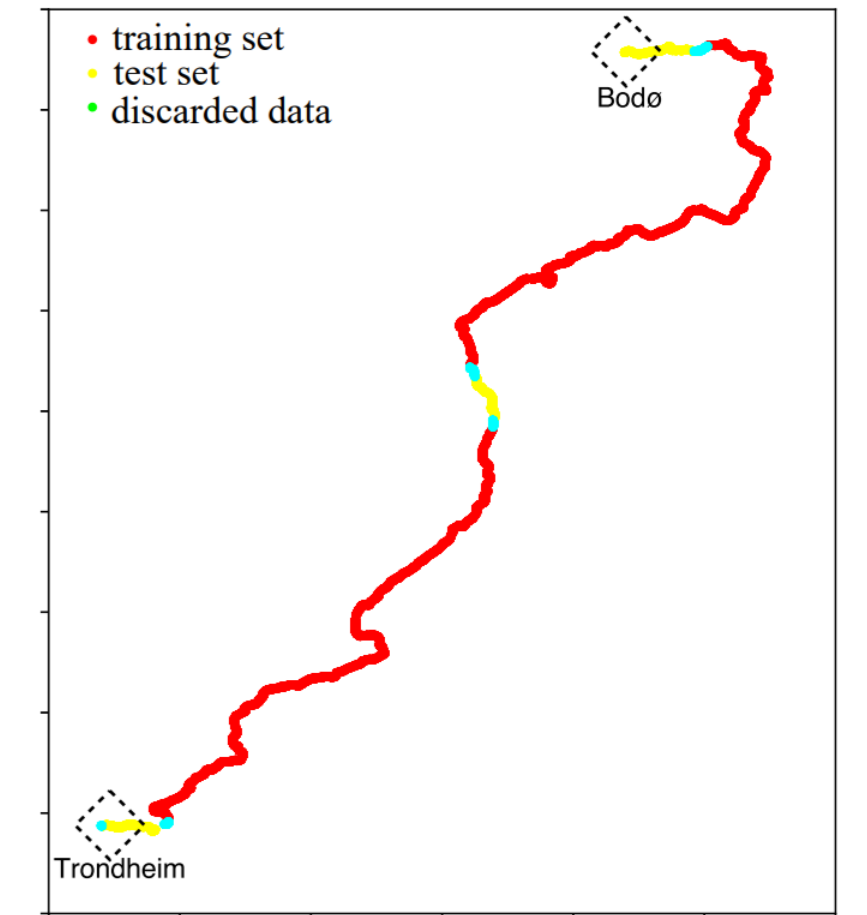
1 Introduction

- Goal: Weather invariant place recognition method based on images.
- Proposal: Siamese and Triplet architectures to extract robust descriptors.



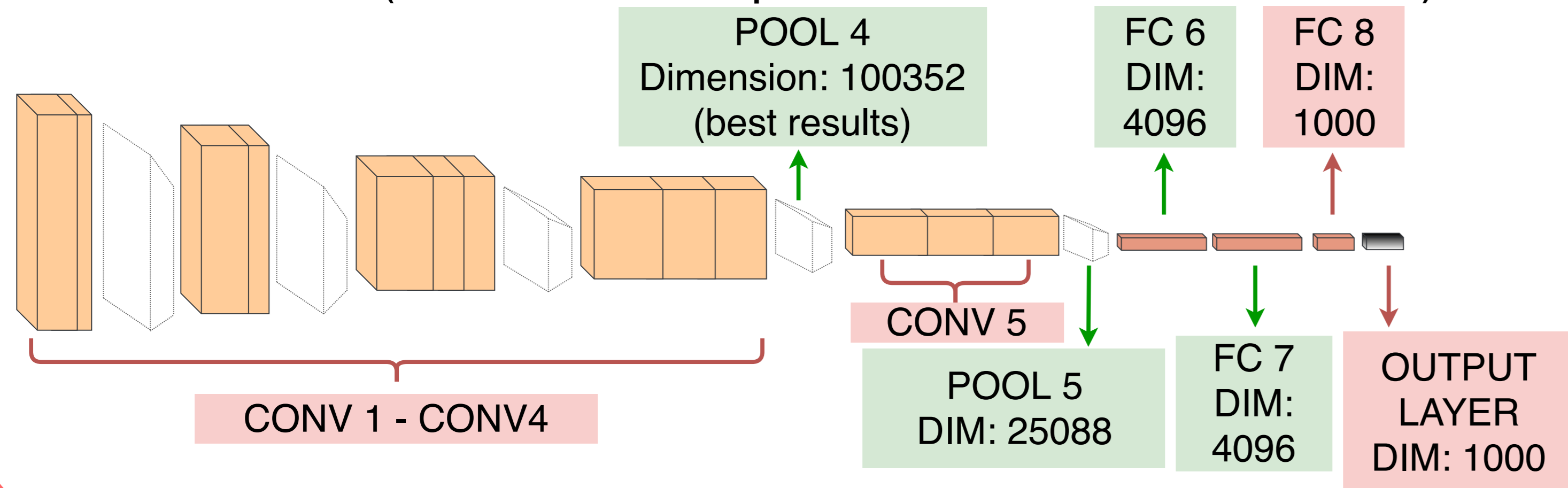
2 Dataset

- Images extracted from Norwegian documentary.
- Used by other researchers with no consensus.
- We propose the following data division:
 - 3,450 Images for testing
 - 24,569 Images for training

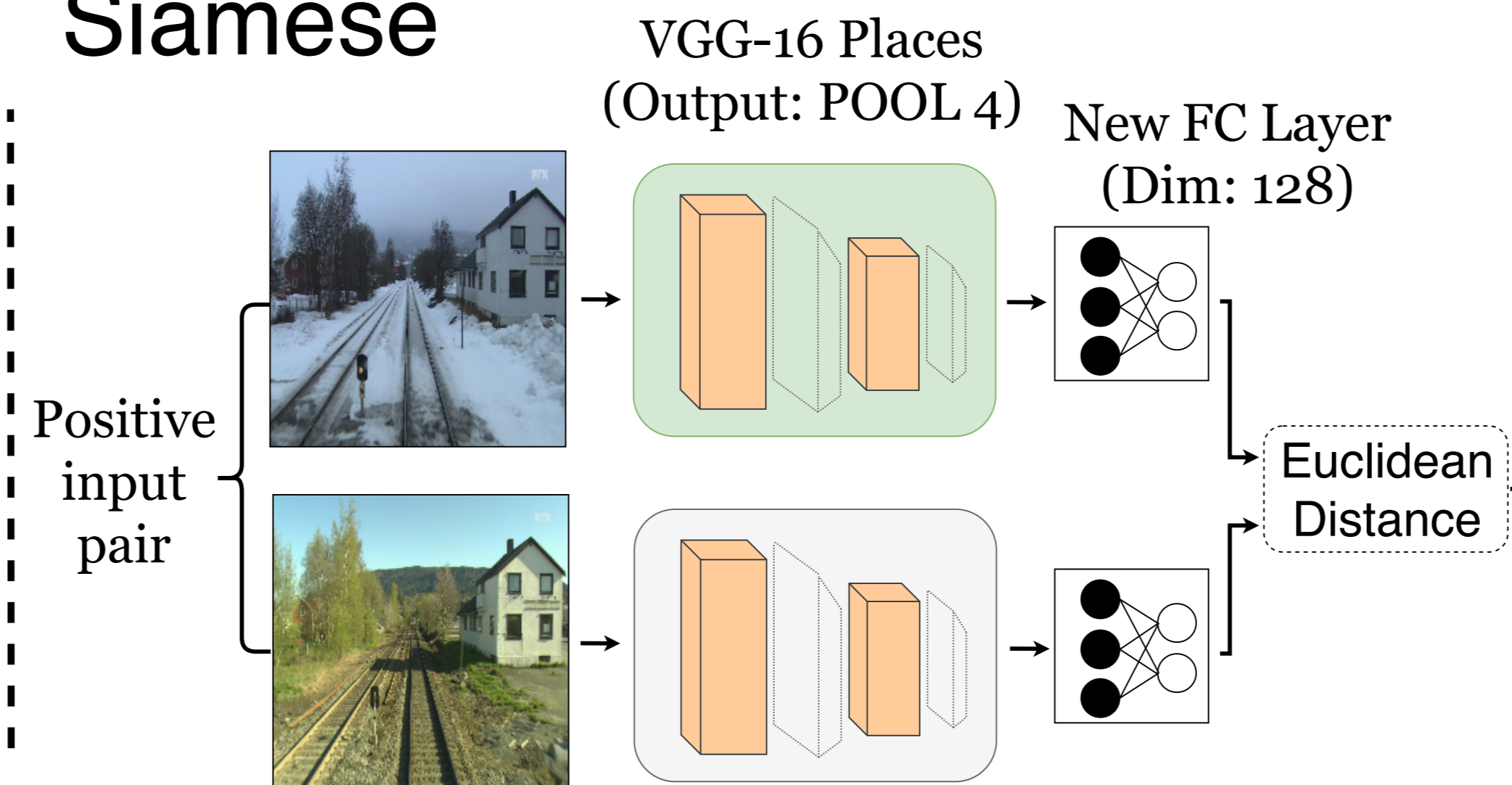


3 Neural Architectures

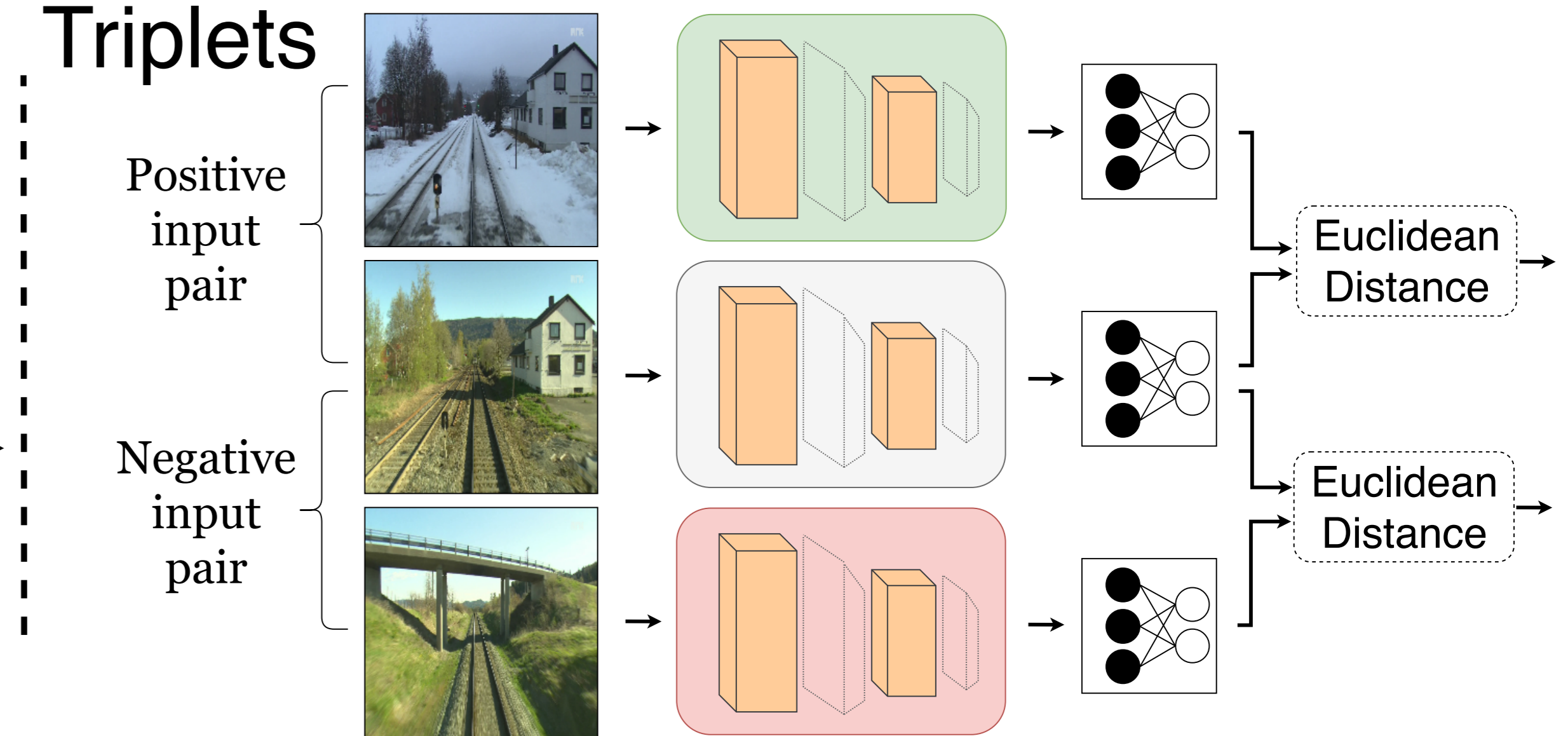
Pre-trained (VGG-16 model pre-trained on Places dataset)



Siamese



Triplets



4 Results

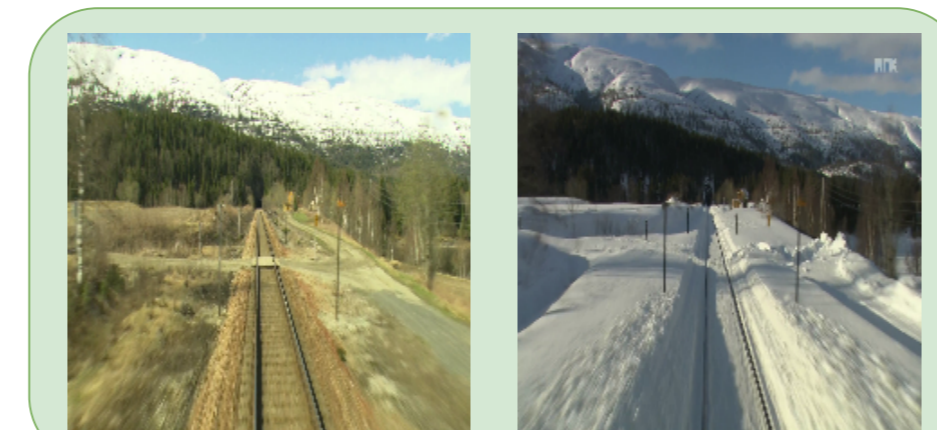
$$\text{Fraction of correct matches} = \frac{\text{N}^\circ \text{ of correctly predicted places}}{\text{N}^\circ \text{ of evaluated places}}$$

NORDLAND		
Architecture	Input: Winter Reference: Summer	Query: Summer Reference: Winter
VGG-16 Pool4	55%	24%
Siamese fc128	60%	61%
Triplet fc128	75%	79%
Triplet fc128 (fine tuned)	86%	86%

ALDERLEY	
Architecture	Query: Day Reference: Night
Triplet fc128 (fine tuned on Nordland dataset)	0,15%
Triplet fc128 (trained on Alderley dataset)	6,84%

SANTA LUCÍA	
Architecture	Mean of all combinations
Triplet fc128 (fine tuned on Nordland dataset)	40%

Same places correctly matched



Different places wrongly matched



5 Conclusions

- Triplet and siamese networks extract features robust to the appearance changes learned during training.
- Best results obtained with a fine tuned triplet structure based on the VGG-16 model pre-trained on the Places dataset.
- State-of-the-art results on the Nordland dataset.



Project Page



ArXiv.org