

# Multi to multi object near-neighbour search

Kamyar Babazadeh, Ben Claydon, Richard Connor, Alan Dearle, Graham Kirby  
{kb267, bc89, rchc, alan.dearle, graham.kirby} @st-andrews.ac.uk

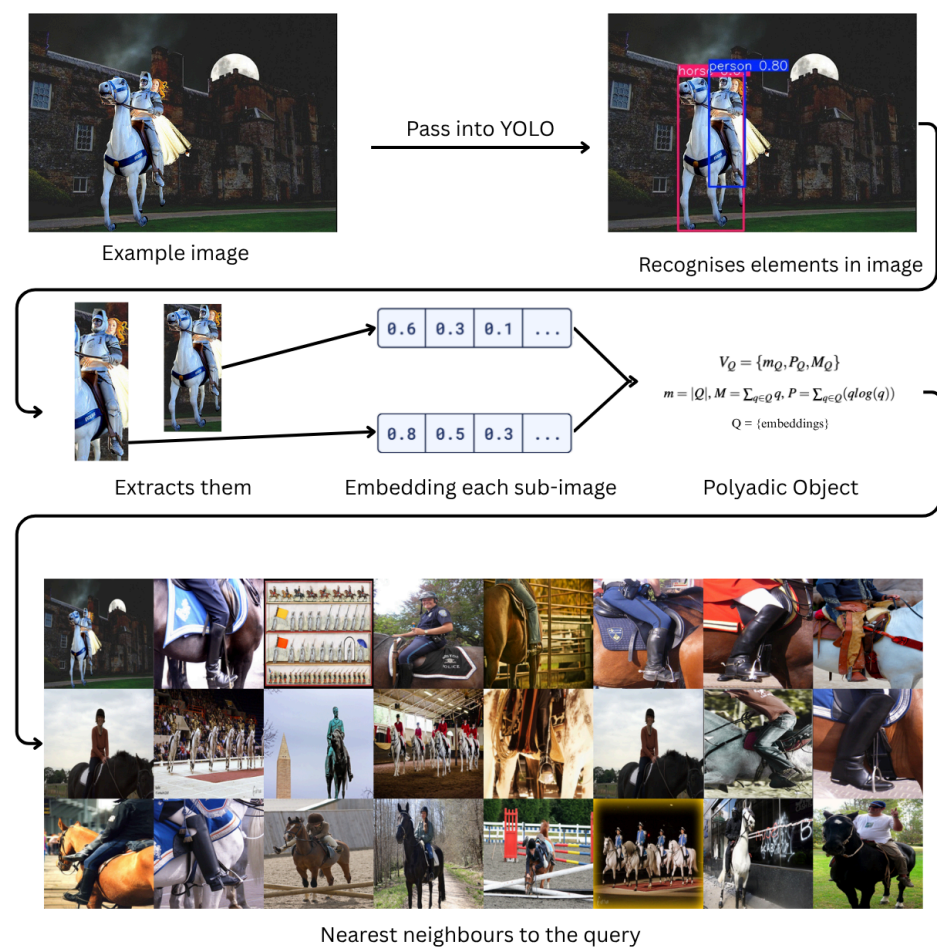
## Introduction

- Near Neighbour Search is the problem of finding the most similar objects in a database to another, presented as a query.
- Recent work [1][2] has demonstrated successful multi query search, where a set of independent objects is used instead to search the collection of singular objects.
- The MSED [3] divergence function is a key part of this work.

## Storing multi-object representations in a database

- We first generate vector embeddings for each object.
- We then softmax them to get discrete sets of probabilities for each object.
- We then store:
  - Entropy of the arithmetic mean of all the objects
  - Geometric mean of their individual entropies
  - Number of objects in the set

## Finding similar images using their sub-images as the query



Versus the nearest neighbours to the example image if we used the whole image



## Searching for the nearest neighbours

To find the nearest neighbour of a multi-object query set we:

1. Add the other object sets to the query set.
2. Calculate the general dissimilarity over all the elements using the MSED divergence of the resulting set.
3. Then rank the search space by object sets that result in smallest-largest divergence when added to the query set.

## Results

- MSED can be effective at performing similarity search over a collection of sets of images in certain scenarios but can also be significantly off in some cases.
- Using images with repeated sub-images return worse outcomes than those with distinct sub-images.

[1] R. Connor, A. Dearle, D. Morrison, and E. Chávez, "Similarity Search with Multiple-Object Queries," Lecture notes in computer science, vol. 14289, pp. 223-237, Jan. 2023, doi: [https://doi.org/10.1007/978-3-031-46994-7\\_19](https://doi.org/10.1007/978-3-031-46994-7_19).

[2] B. Claydon, R. Connor, A. Dearle, and L. Vadicamo, "Demonstrating the Efficacy of Polyadic Queries," Lecture notes in computer science, vol. 15268, pp. 49-56, Oct. 2024, doi: [https://doi.org/10.1007/978-3-031-75823-2\\_4](https://doi.org/10.1007/978-3-031-75823-2_4).

[3] R. Moss and R. Connor, "A Multi-way Divergence Metric for Vector Spaces," Lecture notes in computer science, vol. 8199, pp. 169-174, Jan. 2013, doi: [https://doi.org/10.1007/978-3-642-41062-8\\_17](https://doi.org/10.1007/978-3-642-41062-8_17).