

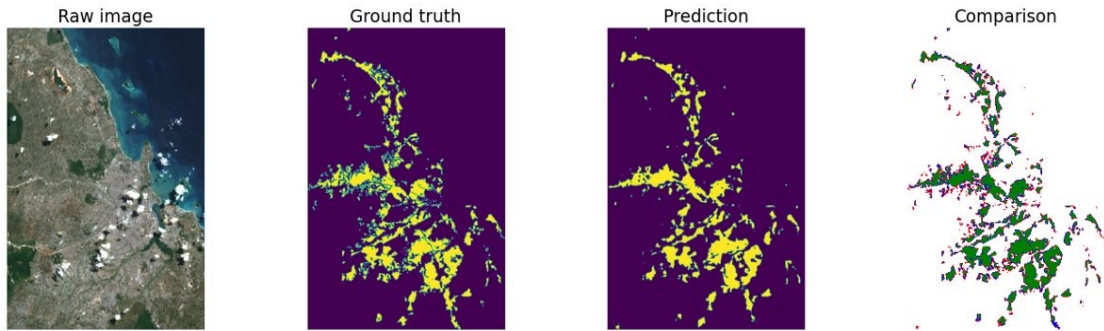
COMPUTER VISION FOR URBAN DEVELOPMENT UNDERSTANDING



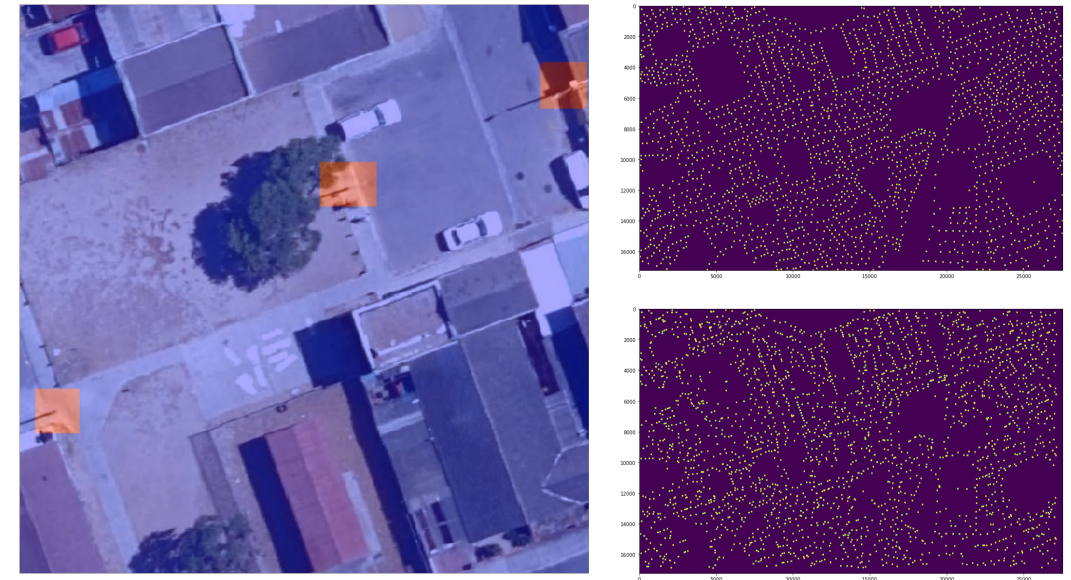
Steven Stalder, **Michele Volpi**, Nicolas Büttner, Esra Suel, Kenneth Harttgen
Swiss Data Science Center and Center for Development and Cooperation, ETH Zürich

Project framework: support development through processing of imagery

Informal settlements (Sub-Saharan Africa)



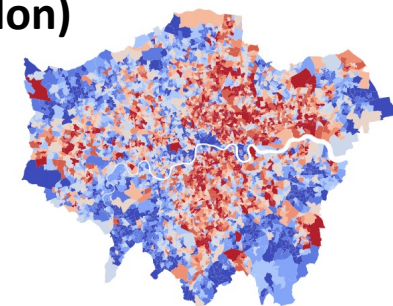
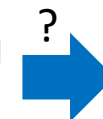
Public lighting estimation, Cape Town



Embedding of street view images (London)

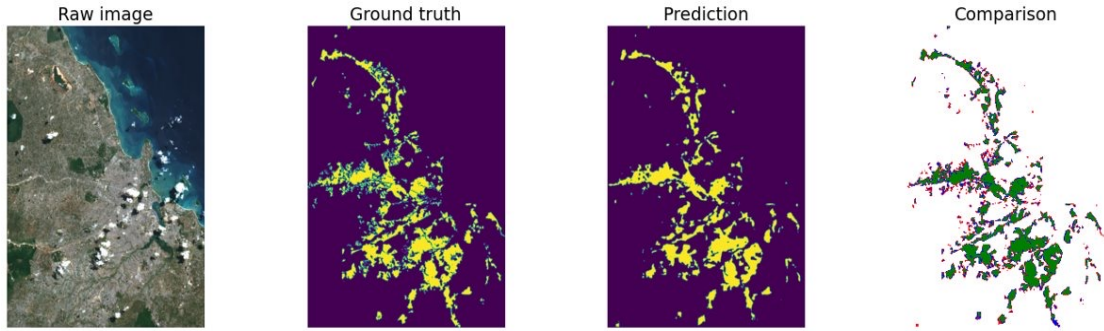


Self-supervised pretraining

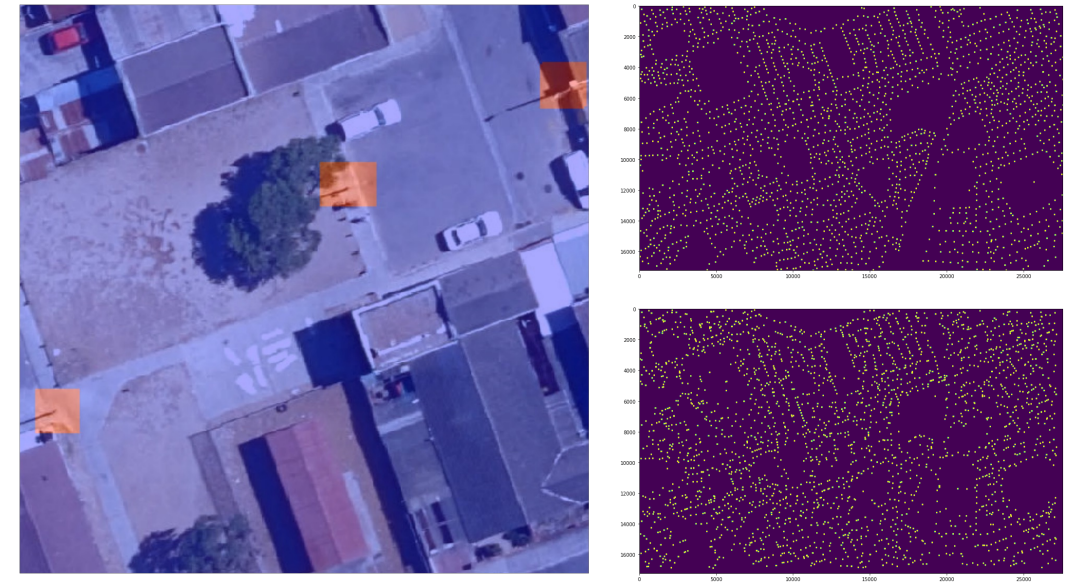


Project framework: support development through processing of imagery

Informal settlements (Sub-Saharan Africa)



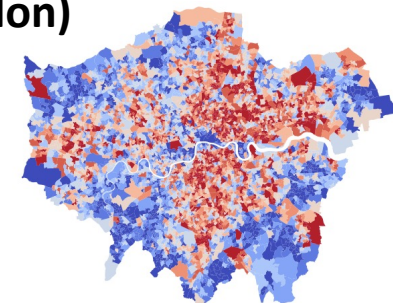
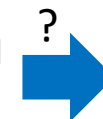
Public lighting estimation, Cape Town



Embedding of street view images (London)

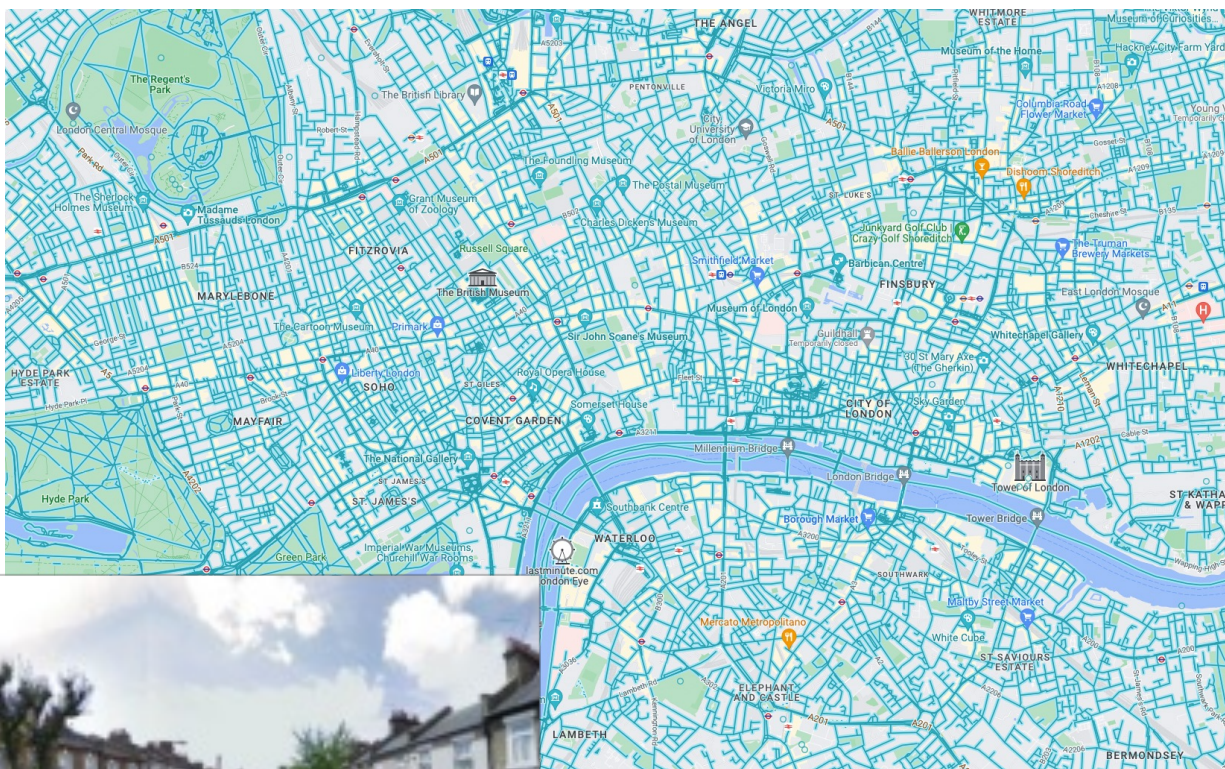


Self-supervised
pretraining



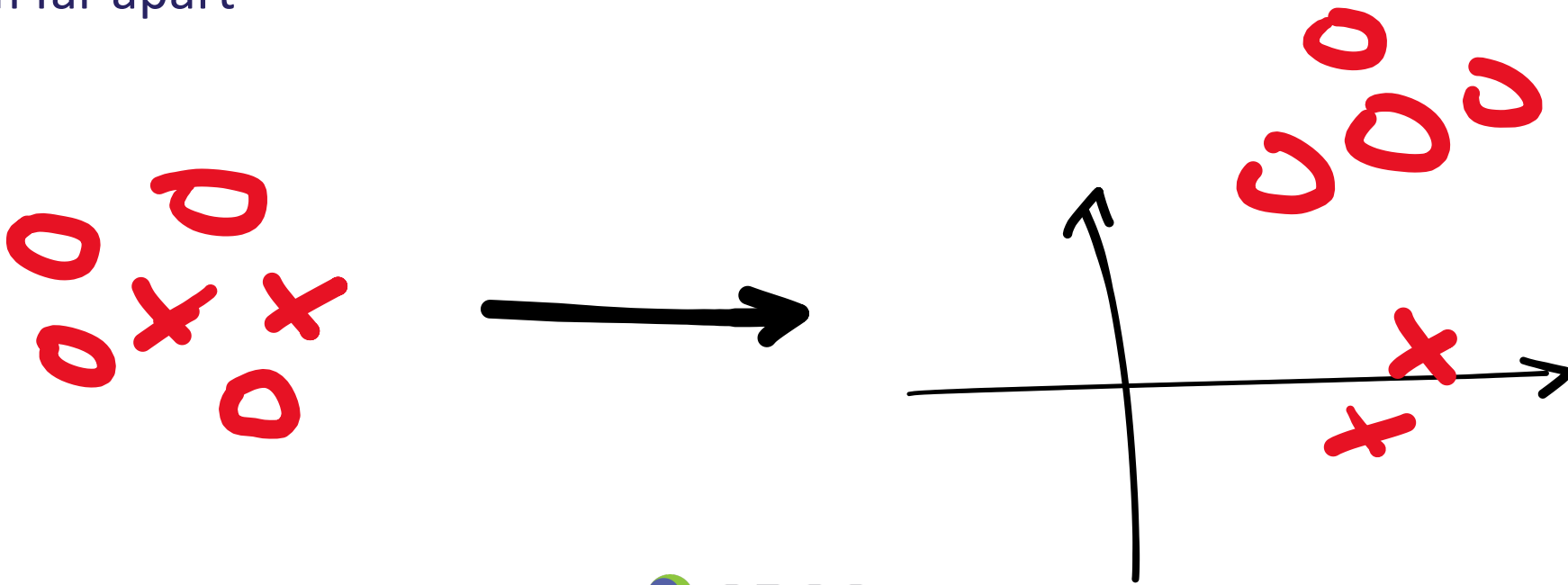
Street view panoramas

- High-level visual detail of urban structures in space and time
- Use this source of visual information for other urban inference tasks
- Need for a simplified, generic data embedding

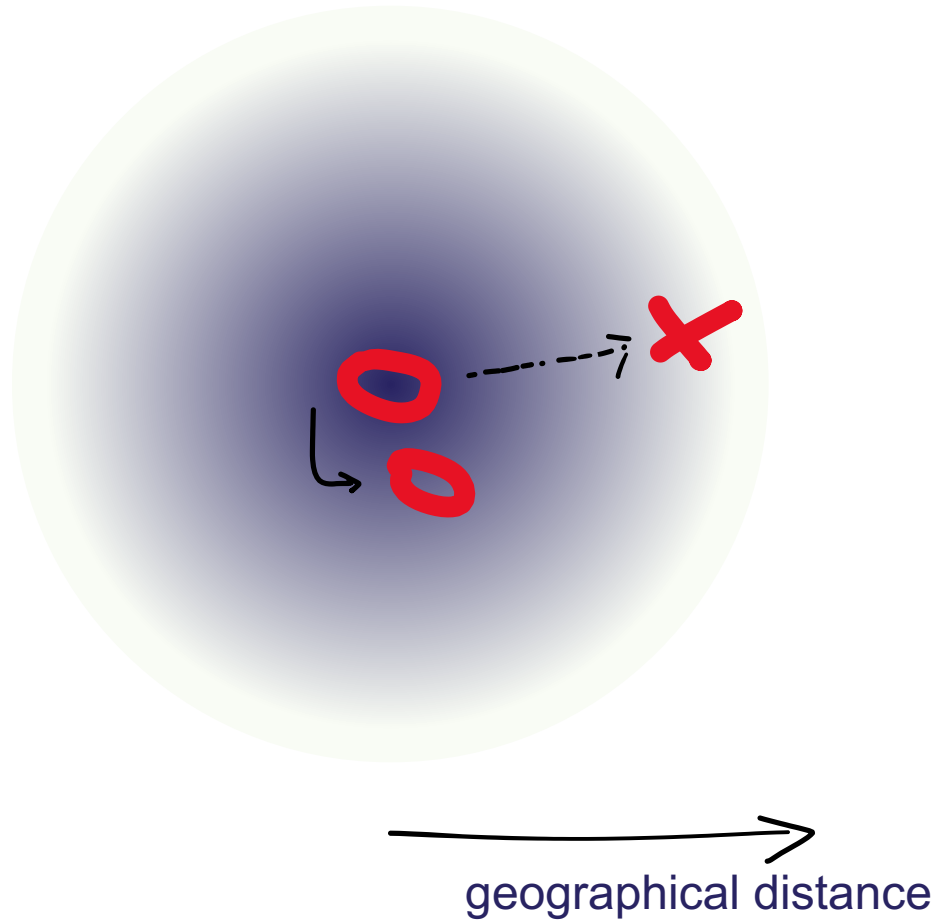


Learn street view panoramas representations

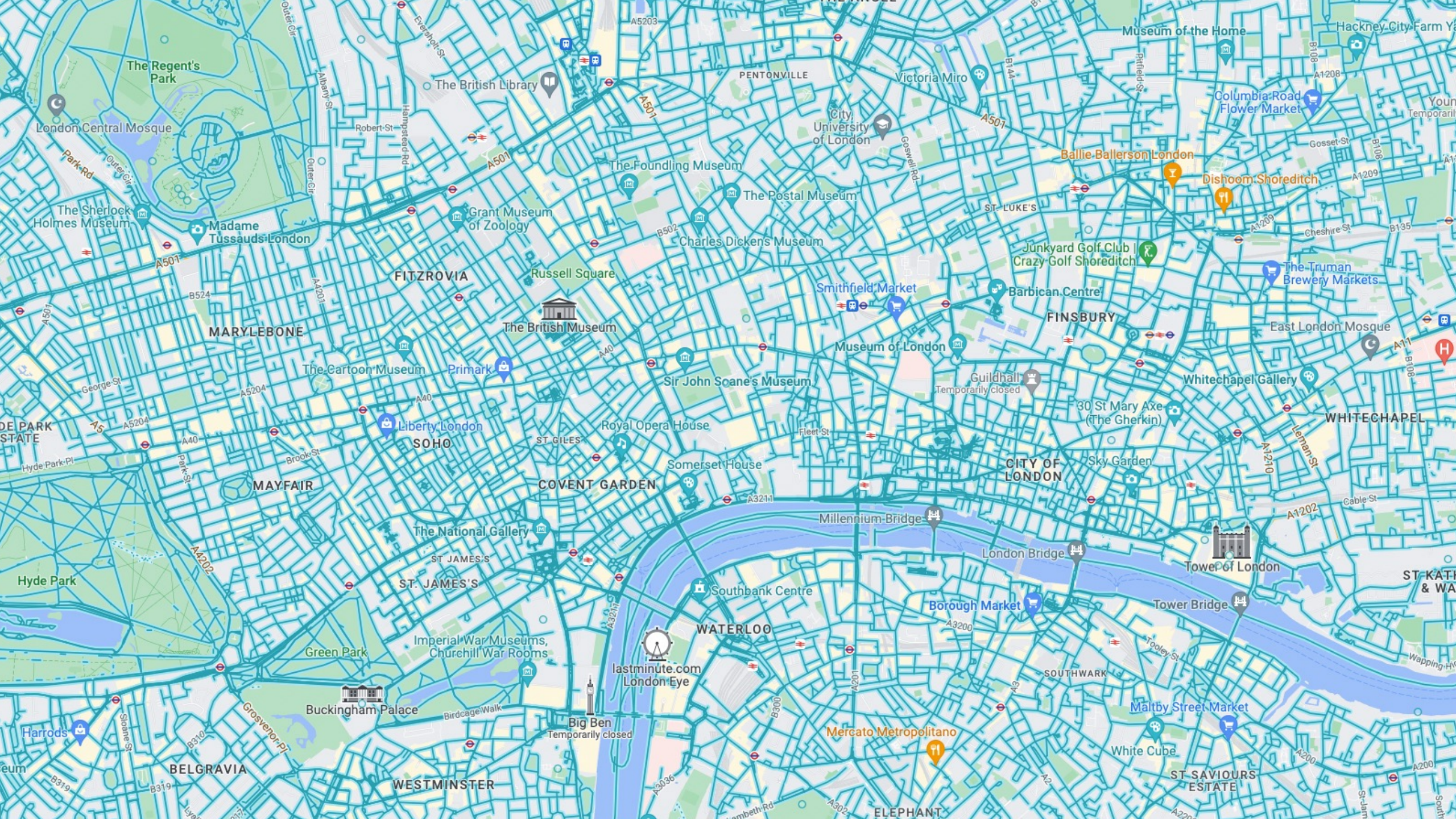
- Contrastive self-supervised learning:
 - Similar objects: learn to map them close-by
 - Dissimilar objects: learn to map them far apart



Learn street view panoramas representations



- What is similar and dissimilar in geographical space?
 - Tobler's first law of geography: "Everything is related to everything else, but near things are more related than distant things"



The Regent's Park

London Central Mosque

The Sherlock Holmes Museum

Madame Tussauds London

FITZROVIA

MARYLEBONE

The Cartoon Museum

Primark

Liberty London

MAYFAIR

Hyde Park

The National Gallery

ST. JAMES'S

Green Park

Buckingham Palace

BELGRAVIA

WESTMINSTER

COVENT GARDEN

The National Gallery

lastminute.com London Eye

Big Ben Temporarily closed

WATERLOO

Mercato Metropolitan

ELEPHANT

PENTONVILLE

The Foundling Museum

The Postal Museum

Charles Dickens Museum

Russell Square

The British Museum

Sir John Soane's Museum

Royal Opera House

Somerset House

Museum of London

Guildhall Temporarily closed

City of London

London Bridge

Borough Market

SOUTHWARK

Maltby Street Market

White Cube

ST SAVIOURS ESTATE

PENTONVILLE

Victoria Miro

City University of London

Ballie Ballerson London

Dishoom Shoreditch

Junkyard Golf Club | Crazy Golf Shoreditch

Barbican Centre

FINSBURY

East London Mosque

Whitechapel Gallery

WHITECHAPEL

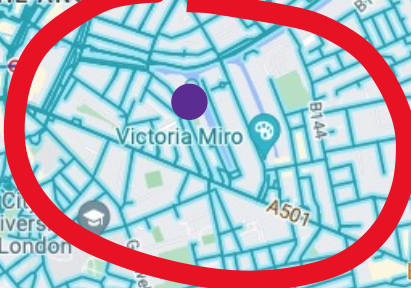
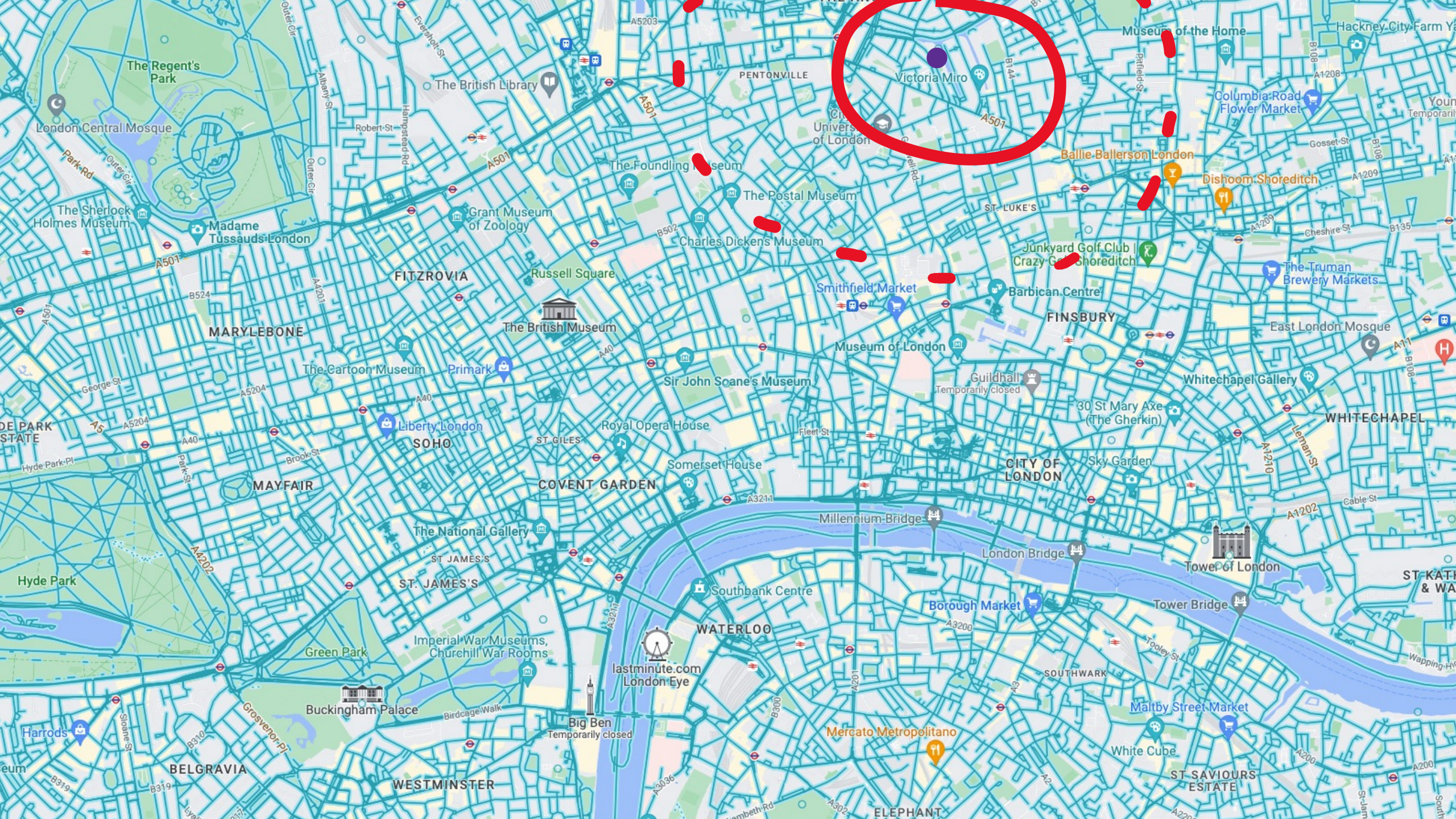
30 St Mary Axe (The Gherkin)

Sky Garden

Tower of London

Tower Bridge

ST KATH & WA



Victoria Miro

PENTONVILLE

The British Museum

Museum of London

Tower of London

lastminute.com London Eye

Buckingham Palace

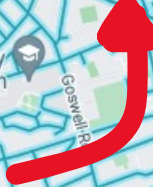
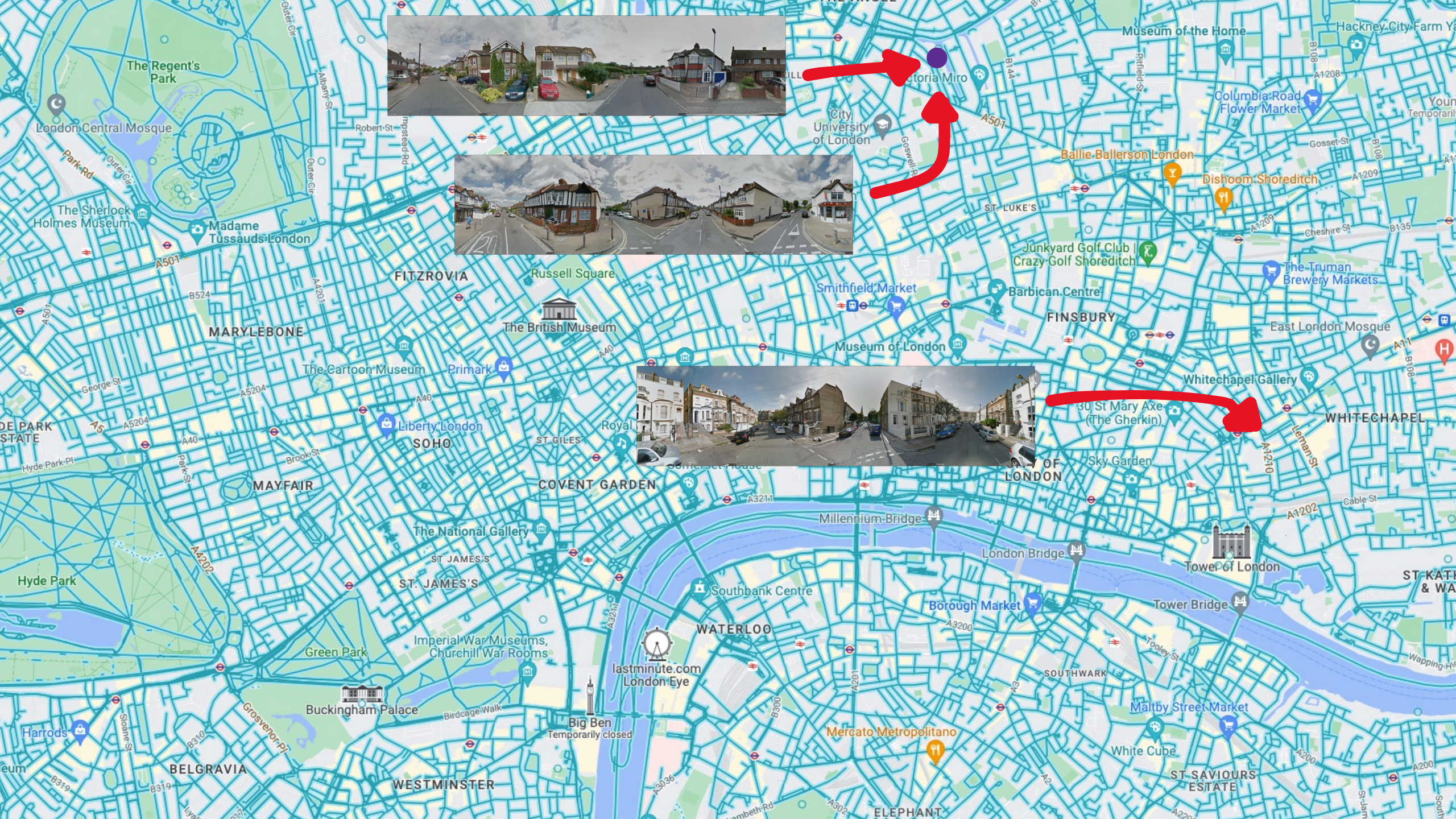
Mercato Metropolitan

BELGRAVIA

WESTMINSTER

ELEPHANT

ST SAVIOURS ESTATE



Anchor (a)



Positive (p)

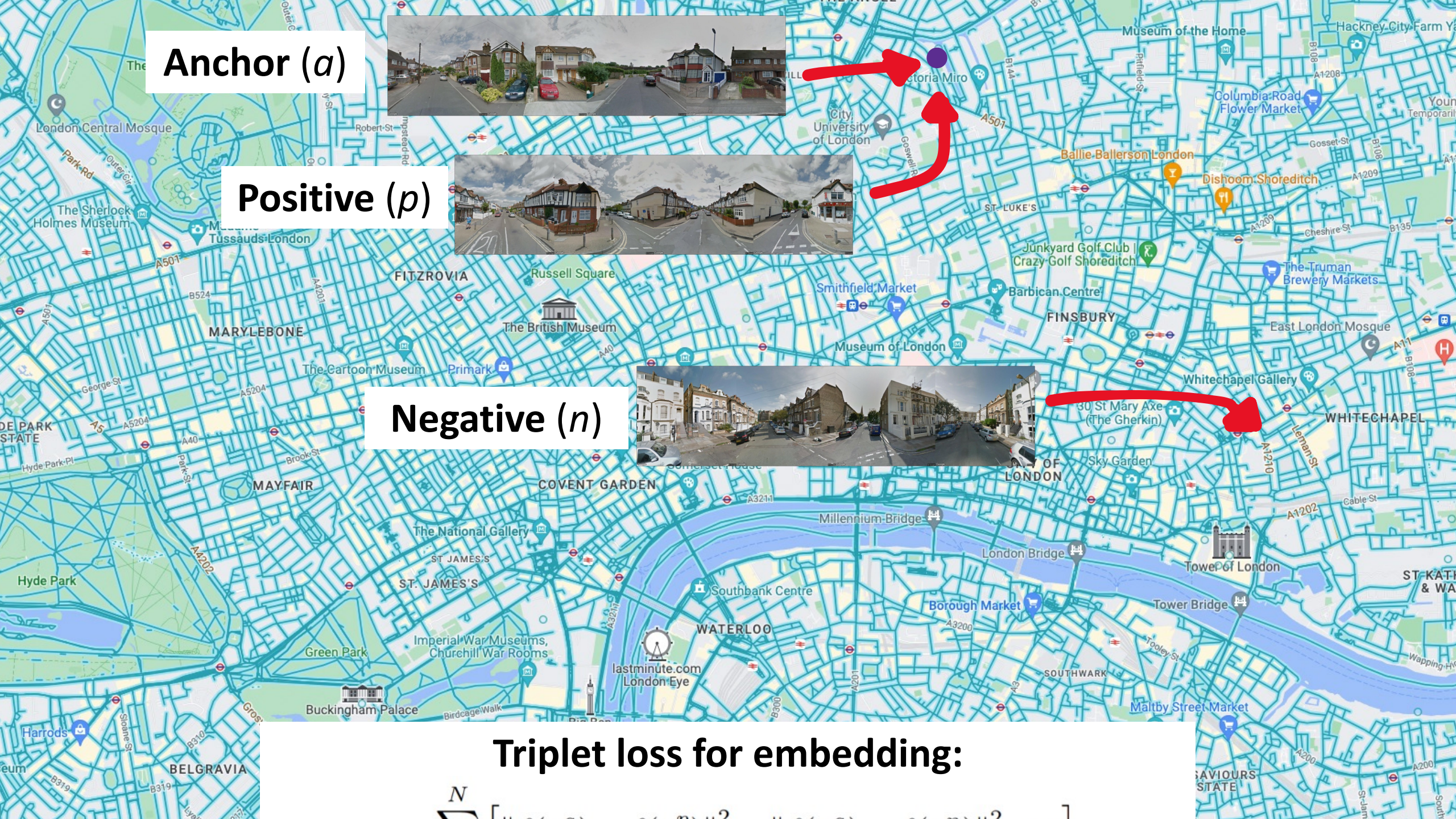


Negative (n)

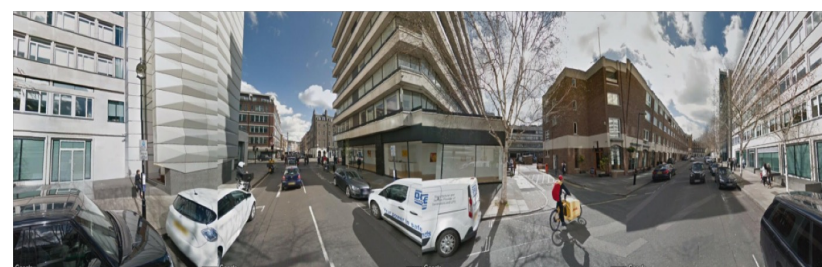


Triplet loss for embedding:

$$\sum_N [\| \mathbf{u}_a - \mathbf{u}_p \| - \| \mathbf{u}_a - \mathbf{u}_n \|]$$



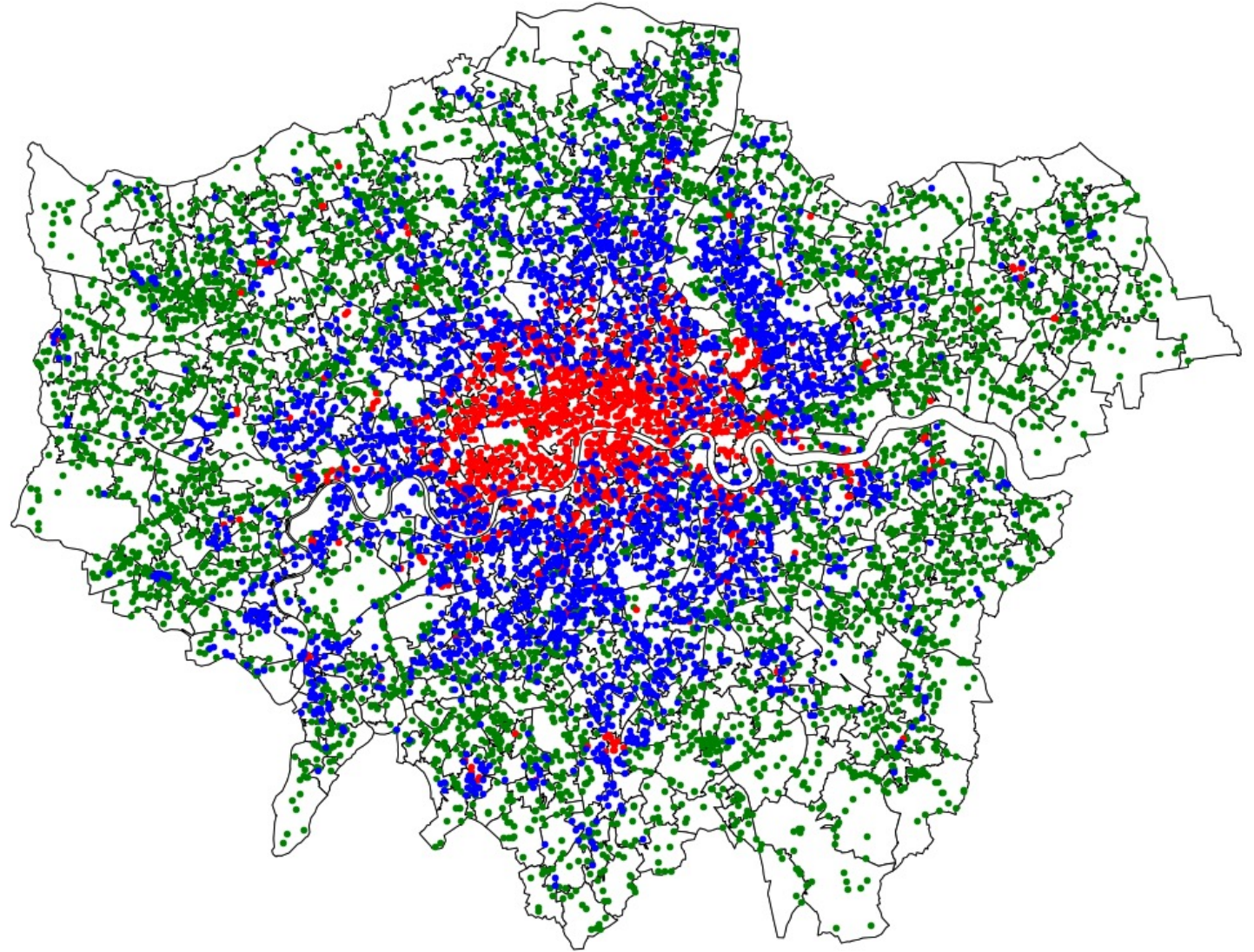
StreetView2Vec - Idea



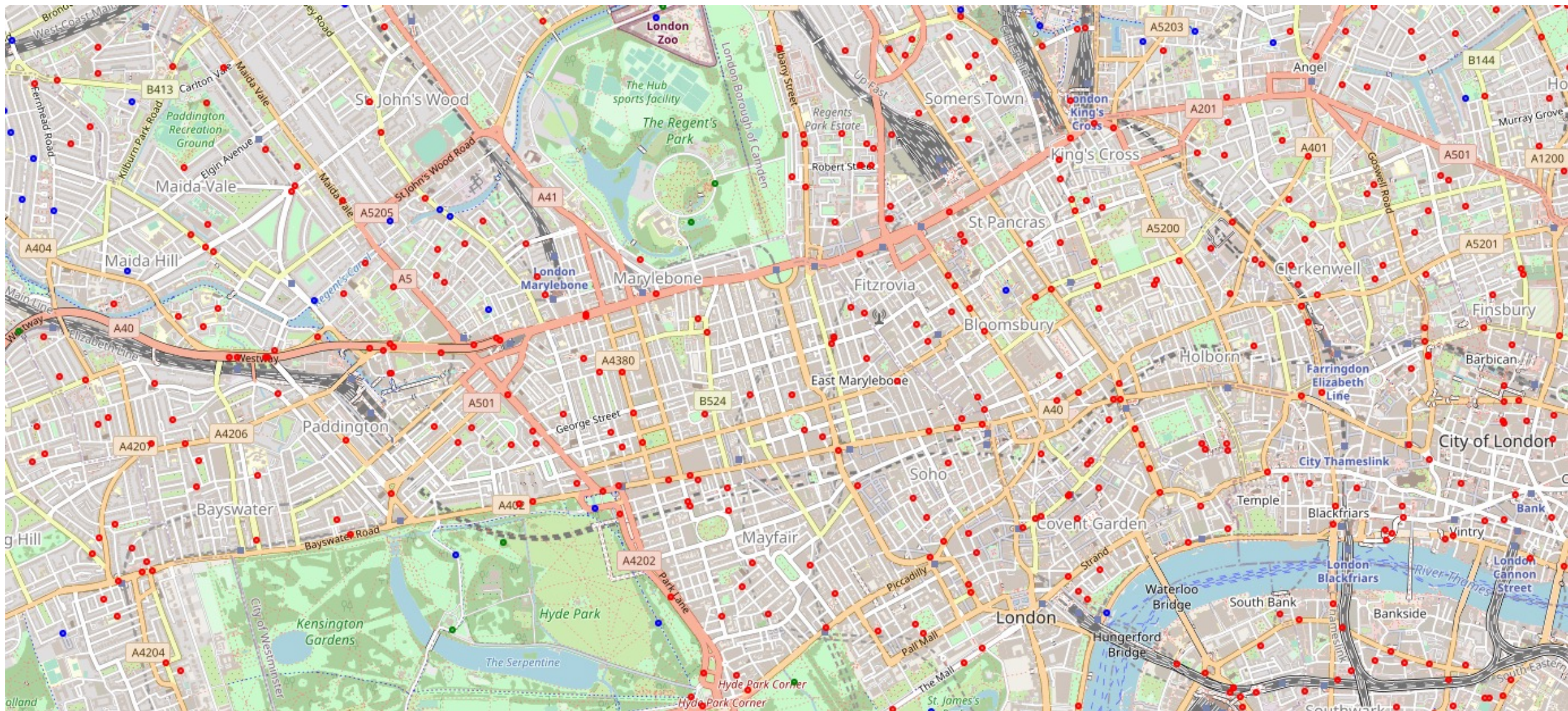
StreetView2Vec – Visualisation of embeddings

- 64-dim embedding learned through Tobler's triplet loss
- Semantics assessed through clustering
 - k-means with 3 clusters
(better visualization using t-SNE or UMAP ongoing work)

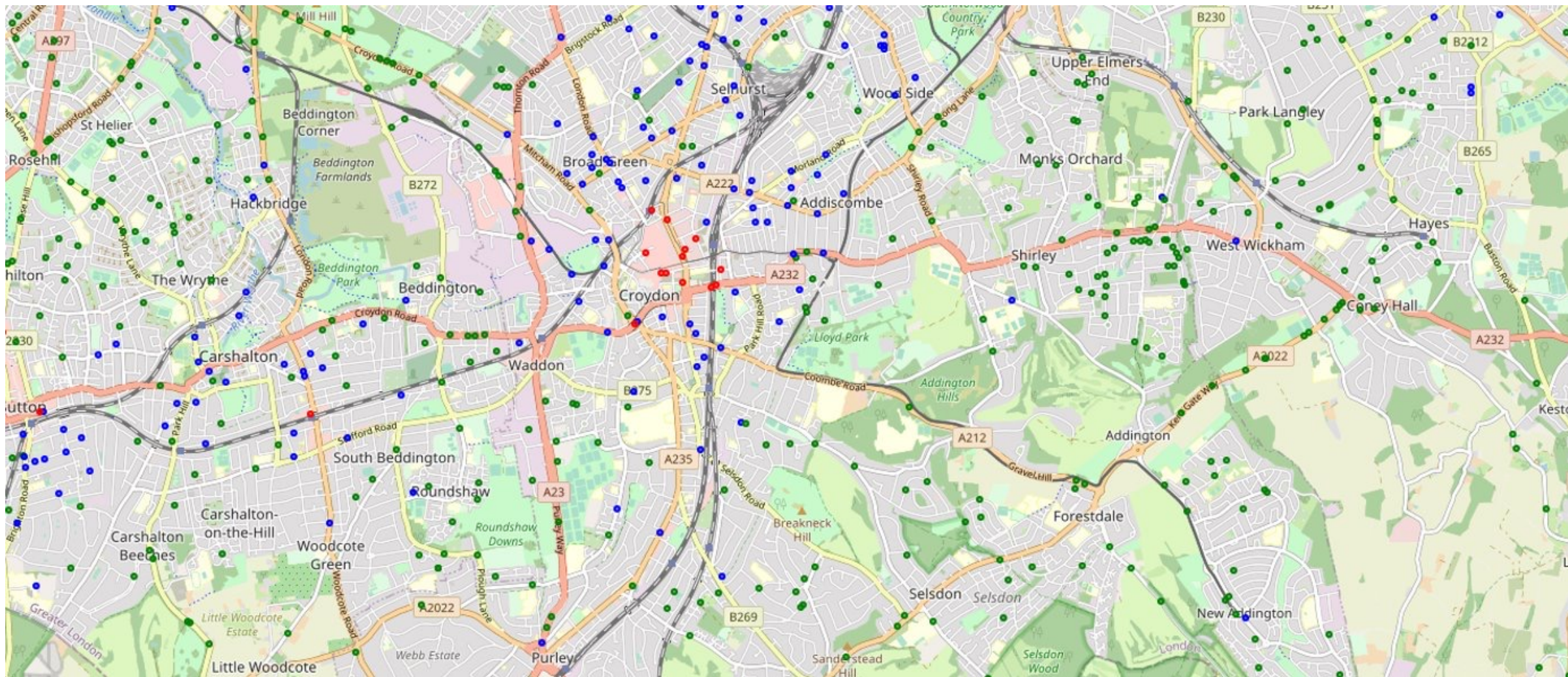
→ Model has **no knowledge** of absolute locations, only knows about relative visual cues!



StreetView2Vec – Some interesting findings



StreetView2Vec – Some interesting findings



StreetView2Vec – Measuring Change

2008



2018



Changed areas




Unchanged areas



ETH zürich

EPFL

PAUL SCHERRER INSTITUT
PSI

 **WSL** Eidg. Forschungsanstalt für Wald
Schnee und Landschaft WSL

 **Empa**
Materials Science and Technology

eawag
aquatic research 000

StreetView2Vec – Measuring Change

2008



2018



StreetView2Vec – Measuring Change

2008



2018



StreetView2Vec – Measuring Change

2008



2018



Learn street view panoramas representations

- Contrastive self-supervised learning:
 - Similar objects: learn to map them close-by
 - Dissimilar objects: learn to map them far apart
- What are we learning?
 - We train a ML model that **maps geolocated images in a simpler, semantically meaningful, latent space.**
 - Other methods could use this **representation better** than raw images
- What is similar and dissimilar in geographical space?
 - Tobler's first law of geography: "Everything is related to everything else, but near things are more related than distant things"