Automated Analysis of Flag Displays in Northern Ireland: A Computer Vision Approach for Cultural and Political Symbolism Recognition

Barry Quinn

2025-03-11

1 Introduction

This research proposal outlines a project that applies advanced computer vision and natural language processing techniques to analyse flag displays in Northern Ireland (NI). The study aims to develop and refine zero-shot object detection and classification methodologies to identify, categorise, and analyse flags in public spaces across NI's urban environments. This work builds upon an existing research foundation that has already demonstrated promising results in detecting flag displays using Google Street View imagery. As Bryan et al. (2010) have established, flags serve as powerful territorial markers in post-conflict Northern Ireland, making their systematic analysis valuable for understanding cultural and political dynamics.

2 Background and Context

Flag displays in Northern Ireland represent significant cultural, political, and community identifiers with substantial implications for understanding social dynamics, territorial marking, and community relations (Jarman 2005; Bryan 2015). Traditional methods of documenting and analysing these displays have been limited by resource constraints, geographical coverage, and temporal inconsistencies. The application of artificial intelligence, specifically computer vision algorithms, offers an innovative approach to systematically capture and analyse these visual symbols at scale.

The project leverages an extensive dataset of approximately two million Google Street View images collected during 2022-2023 across Northern Ireland's 50 largest urban areas. Initial analysis using GroundingDINO (Liu et al. 2023), a zero-shot object detection model, has

identified approximately 100,000 potential flag instances, with manual verification confirming approximately 70,000 true positives.

3 Research Aims and Objectives

This research aims to develop and implement advanced AI methodologies for the automated detection, classification, and analysis of flag displays in Northern Ireland. The specific objectives include:

1. Refine the existing GroundingDINO-based detection system through further training with verified examples to improve precision whilst maintaining zero-shot flexibility
2. Develop a robust classification framework using vision-language models to categorise detected flags into seven primary categories (National, Fraternal, Sport, Military, Historical, International, and others)
3. Design and implement a hierarchical prompt tuning methodology informed by recent advancements in fine-grained visual classification (Li et al. 2023)
4. Establish a workflow for the systematic analysis of flag display patterns in relation to geographical, temporal, and socio-economic factors
5. Create a transferable methodological framework that can be applied to historical imagery and potentially other cultural contexts

4 Methodology

4.1 Detection Refinement

The research will build upon the existing implementation of GroundingDINO (Liu et al. 2023), which combines a vision transformer architecture with a BERT-based text encoder. This unique architecture enables zero-shot object detection capabilities without requiring specialised training data. However, to enhance detection accuracy for the specific task of flag identification, the model will be refined using the manually verified dataset of true and false positives already collected. This approach aligns with the methodology proposed by Kirillov et al. (2023) for adapting foundational models to specific detection tasks.

4.2 Classification Framework Development

Drawing inspiration from semantic-assisted object clustering techniques (Luo et al. 2023), this project will implement a hierarchical classification system. The classification framework will incorporate:

1. A structured labelling process involving domain experts to classify 3,000-5,000 strategically sampled images
2. Implementation of a custom web interface for hierarchical classification with confidence scoring
3. Adaptation of prompt tuning methods similar to those applied in fine-grained ship classification by Li et al. (2023) in their seminal work on vision-language models for object classification

The approach builds on Li et al.'s (2023) "Semantic-Assisted Object Clustering for Referring Video Object Segmentation," which demonstrated the effectiveness of integrating semantic information for improved object identification across varied contexts.

4.2.1 Existing Labelling Infrastructure

A critical component of the proposed methodology has already been implemented, significantly enhancing the project's viability. A Next.js-based labelling application has been developed and deployed (accessible at https://flag-labeller-19wrdqvle-barry-quinns-projects.vercel.app/), providing a specialized interface for domain experts to efficiently categorise flag imagery. This application incorporates:

1. Secure authentication protocols to ensure data integrity
2. A hierarchical classification system aligned with the project's taxonomic requirements
3. An intuitive interface designed specifically for domain experts without technical background
4. Data collection mechanisms that include confidence metrics and inter-annotator reliability measures

Three domain experts have already been identified and committed to completing the labelling of a strategically sampled dataset by summer 2025. This existing infrastructure and confirmed expert participation substantially reduces implementation risks and ensures that the critical training dataset will be available within the proposed timeline.

4.3 Data Processing and Analysis

The methodology will include:

1. Preprocessing of Google Street View imagery for optimal model ingestion
2. Implementation of batch processing techniques to handle the large-scale dataset
3. Development of spatial analysis tools to map flag distributions against geographical and administrative boundaries
4. Integration with socio-economic indicators to enable multidimensional analysis of flag display patterns

This approach draws on methodologies established by Gebru et al. (2017) for using computer vision to analyse demographic patterns through Google Street View imagery.

5 Technical Implementation

The project will employ state-of-the-art deep learning frameworks including:

1. Vision Transformers for feature extraction from imagery (Dosovitskiy et al. 2021)
2. Vision-Language Models for semantic understanding and context awareness (Radford et al. 2021)
3. Hierarchical classification architectures to capture nuanced flag typologies
4. Geospatial data processing tools for location-based analysis

The implementation will build upon existing Python-based frameworks for computer vision and will be designed with scalability and reproducibility as core principles, following best practices established by Li and colleagues (2023) in their work on semantic-assisted classification.

6 Expected Outcomes and Contributions

This research is expected to deliver:

1. A refined methodology for zero-shot flag detection with improved accuracy metrics
2. A novel hierarchical classification system for flag typology recognition
3. A comprehensive dataset of classified flag displays across Northern Ireland
4. Analytical insights into the spatial and temporal patterns of flag displays
5. A transferable technical framework for analysis of visual cultural symbols in public spaces

The project will contribute to both computational methods in computer vision and to the empirical understanding of visual political symbolism in divided societies (Bryan 2015; Komarova and Torney 2020).

7 Timeline and Resources

The MSc dissertation project will be conducted over a 12-week period, structured as follows:

- Weeks 1-2: Literature review and technical familiarisation
- Weeks 3-4: Development of classification framework
- Weeks 5-7: Implementation of hierarchical prompt tuning and model training
- Weeks 8-9: Evaluation and refinement of classification performance
- Weeks 10-11: Spatial analysis and integration with socio-economic indicators

- Week 12: Final analysis, documentation, and thesis preparation

The project will utilise existing computational resources and datasets already collected by the research group.

8 Ethical Considerations

The research will adhere to best practices in data ethics, ensuring that:

1. All imagery used is publicly available through Google Street View
2. No personally identifiable information is processed or stored
3. Analysis focuses on aggregate patterns rather than specific locations that might identify individual properties
4. Interpretations of findings remain sensitive to the complex cultural and political context of Northern Ireland

This approach follows ethical guidelines established by Crawford and Paglen (2019) for responsible AI research involving public imagery.

9 Supervision and Collaboration

The project will be supervised by Dr Shuyan Li from Computer Science, with additional expertise provided by the interdisciplinary research team that includes specialists in computer vision, anthropology, and Northern Ireland studies. The work builds upon an existing PhD project funded through CAST in collaboration with the Department of Communities and Rural Affairs, ensuring alignment with broader research and policy objectives.

9.1 Project Viability Assessment

This proposal represents an exceptionally viable MSc dissertation project due to several factors that mitigate typical implementation risks:

1. **Existing Dataset**: The two million image dataset has already been collected and processed through initial detection, providing a solid foundation for the classification work.
2. **Operational Infrastructure**: A dedicated Next.js labelling application has been fully implemented and deployed, eliminating development time that would otherwise be required for creating data annotation tools.
3. **Expert Resource Commitment**: Three domain experts have formally committed to completing the annotation of 3,000-5,000 training examples by summer 2025, ensuring the availability of high-quality labelled data at the commencement of the project.

4. **Technical Foundation**: The GroundingDINO implementation has been successfully adapted to the flag detection task, allowing the project to focus on classification refinement rather than fundamental detection challenges.

5. **Interdisciplinary Support**: The established collaboration between Computer Science (Dr Li),Anthropology (Prof Bryan), Health Economics (Prof French) and existing PhD researchers provides a robust support structure for addressing both technical and domain-specific challenges.

These elements collectively ensure that the project can realistically achieve its objectives within the 12-week MSc dissertation timeframe, allowing the researcher to focus on the novel aspects of hierarchical classification and prompt tuning rather than preliminary data collection and infrastructure development.

10 Conclusion

This MSc dissertation project represents a significant opportunity to advance both computational methods in AI and empirical understanding of visual political symbolism. By combining state-of-the-art computer vision techniques with domain expertise in Northern Ireland's cultural landscape, the research aims to develop innovative approaches to the automated analysis of flag displays with potential applications for both academic research and policy development.

References

Bryan, Dominic. 2015. "Parades, Flags, Carnivals, and Riots: Public Space, Contestation, and Transformation in Northern Ireland." *Peace and Conflict: Journal of Peace Psychology* 21 (4): 565–73.

Bryan, Dominic, Clifford Stevenson, Gordon Gillespie, and John Bell. 2010. "Public Displays of Flags and Emblems in Northern Ireland: Survey 2006-2009." Institute of Irish Studies, Queen's University Belfast.

Crawford, Kate, and Trevor Paglen. 2019. "Excavating AI: The Politics of Images in Machine Learning Training Sets." *AI & Society*.

Dosovitskiy, Alexey, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, et al. 2021. "An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale." In *International Conference on Learning Representations (ICLR)*.

Gebru, Timnit, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, and Li Fei-Fei. 2017. "Using Deep Learning and Google Street View to Estimate the Demographic Makeup of Neighborhoods Across the United States." *Proceedings of the National Academy of Sciences* 114 (50): 13108–13.

Jarman, Neil. 2005. "Painting Landscapes: The Place of Murals in the Symbolic Construction of Urban Space." In *Symbols in Northern Ireland*, 81–98. Institute of Irish Studies, Queen's University of Belfast.

Kirillov, Alexander, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, et al. 2023. "Segment Anything." In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*.

Komarova, Milena, and Julie Torney. 2020. "Neutrality in Public Space: A Case Study of Shared Residential Space in Belfast." *Urban Studies* 57 (10): 2032–47.

Li, Shuyan, Zhuoyan Luo, Yicheng Xiao, Yong Liu, Yitong Wang, Yansong Tang, Xiu Li, and Yujiu Yang. 2023. "Semantic-Assisted Object Cluster for Referring Video Object Segmentation." In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*.

Liu, Shilong, Zhaoyang Zhu, Jie Zhang, Feng Zhang, Melih Abdulkadir Davari, Zhenda Chen, Hao Li, et al. 2023. "Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection." *arXiv Preprint arXiv:2303.05499*.

Luo, Zhuoyan, Yicheng Xiao, Yong Liu, Shuyan Li, Yitong Wang, Yansong Tang, Xiu Li, and Yujiu Yang. 2023. "SOC: Semantic-Assisted Object Cluster for Referring Video Object Segmentation." In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*.

Radford, Alec, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, et al. 2021. "Learning Transferable Visual Models from Natural Language Supervision." In *International Conference on Machine Learning (ICML)*.