Artificial Intelligence as the Histopathologist's Assistant: Addressing the Problem of Domain Shift

Rebecca Nachiappan. College of Medical and Dental Sciences, University of Birmingham.

Introduction

- Mitosis detection is important for grading cancers.
- Computational scientists have been developing machine-learning algorithms to automate this process.
- However, the use of different scanners and staining protocols in different centres causes domain shift[3].
- This complicates the development of a model that can be generalised across different centres[3].

Aim

To conduct a literature review on methods used to address domain shift in training a machine learning model for mitosis detection, with a focus on stain normalisation and data augmentation techniques.

Method

- A PubMed search using the search terms "domain shift" and "mitosis detection".
- Search was limited to articles published over the last five years.
- Relevant references within these articles were also included.
- This generated a list of 21 articles for the literature review[1-21].
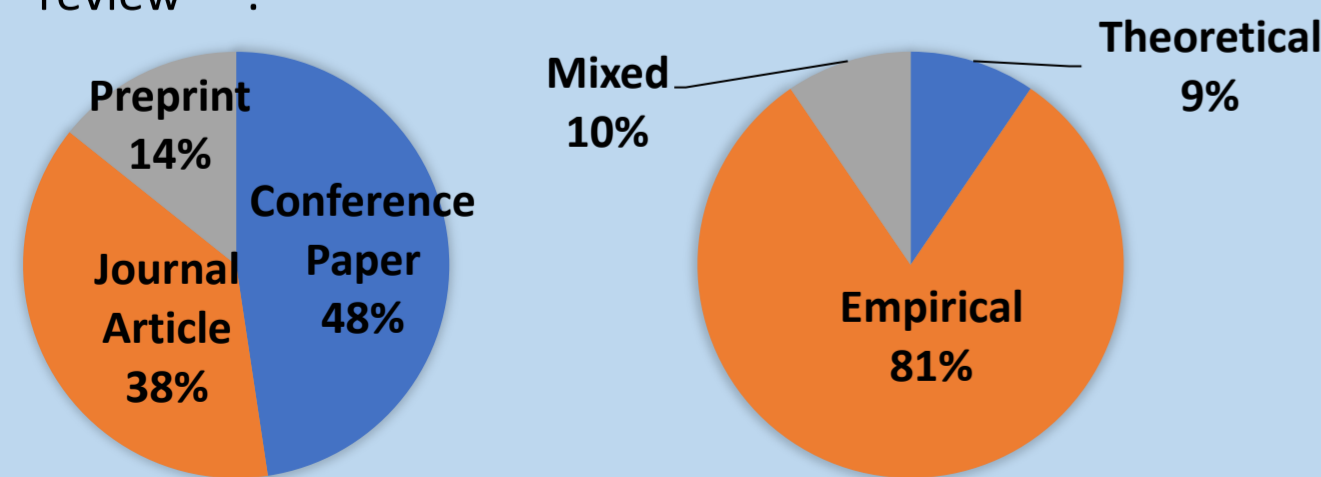

Figure 1: Types of articles in the review.

Figure 2: Types of evidence in the review.

- Stain normalisation was mentioned in 12 articles, while data augmentation was mentioned in 20 articles.
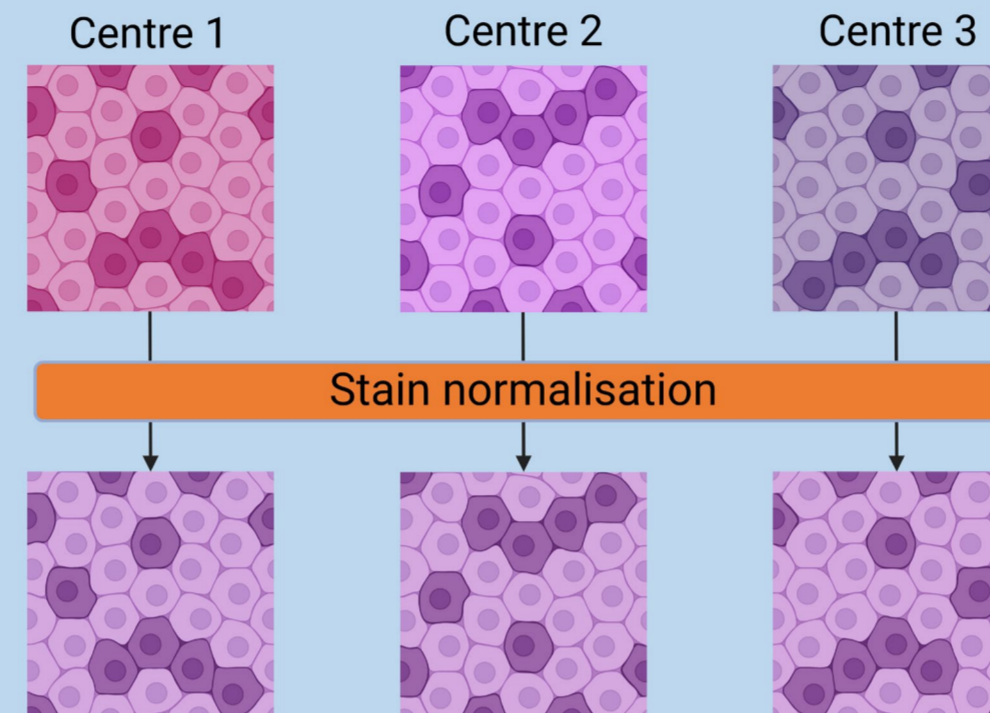
Findings

Figure 3: Illustration of how stain normalisation takes histology slide images from different centres and processes them according to a template using a colour matrix[10]. Created in BioRender. Nachiappan, R. (2024) BioRender.com/l59i291.
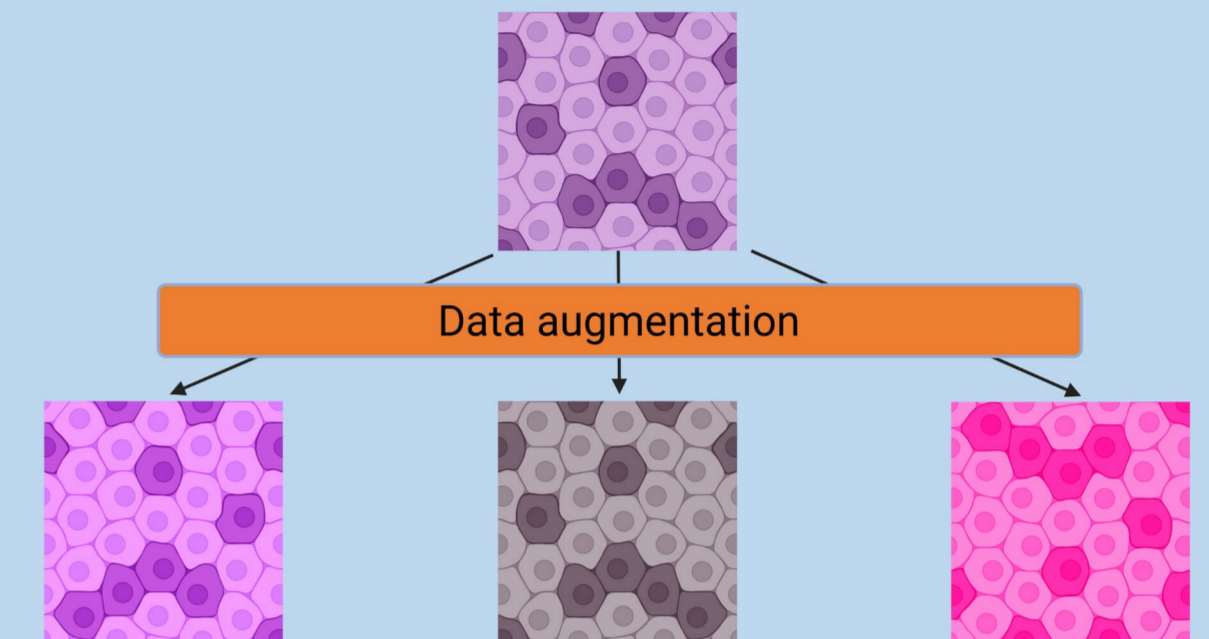
Figure 4: Illustration of how data augmentation takes existing slide images and performs random transformations on them such as flipping, rotating and colour/contrast changes to create synthetic images[10]. Created in BioRender. Nachiappan, R. (2024) BioRender.com/l59i291.

SN alone can be detrimental[1,10,11]

- Does not remove domain-specific differences[15]
- Time-consuming as it adds a lot of computational overhead[15]
- Neural-network methods of SN seem better than traditional methods[10]
- Combining SN with DA methods yields better results[10]

DA outperforms SN in improving model generalisability[4,8,9,12,16,21]

- DA simulates stain variation by creating synthetic images[16]
- DA does not add as much computational overhead[15]
- DA avoids loss of relevant information due to limitations of SN methods[16]

Figure 5: Key points gleaned from the literature review. 'SN' refers to stain normalisation and 'DA' refers to data augmentation.

Conclusion

Stain normalisation and data augmentation are both techniques that can reduce domain shift when developing a machine learning algorithm for mitosis detection. However, data augmentation methods have their strengths over using stain normalisation methods on their own. If only stain normalisation can be used, neural-network based methods are preferable. The research into creating a machine learning model for mitosis detection that can tolerate images with a high degree of variation is still in its infancy.

References