Multi-view analysis of unregistered medical images using cross-view transformers

Gijs van Tulder, Yao Tong, Elena Marchiori

Motivation

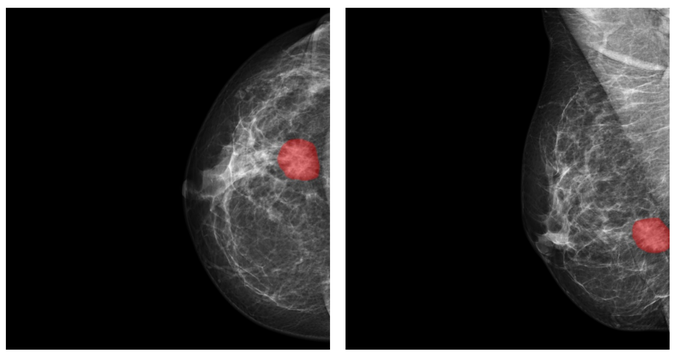
Multi-view medical image analysis often depends on the combination of informaton from multiple views.

Differences in perspective and other forms of misalignment can make it difficult to combine views effectively.

Data

CBIS-DDSM mammography images with CC and MLO views.

Breast tumor classification benefits from information from both views.



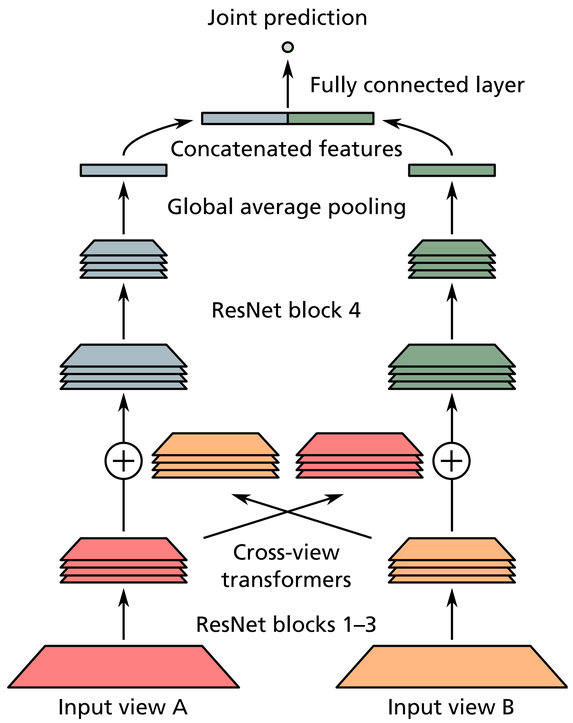
Craniocaudal (CC) *Mediolateral-oblique (MLO)*

Method

A cross-view transformer that can transfer features between unregistered views at the level of spatial feature maps.

Cross-view transformer

Link views and transfer features at an intermediate, spatial level.



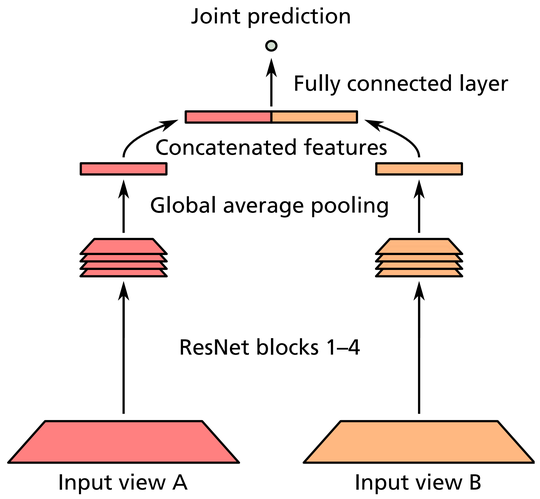
Results for CBIS-DDSM breast tumor classification

Model	Views	ROC-AUC ± std.dev.	p-value
Single view	CC	0.750 ± 0.007	0.005
	MLO	0.763 ± 0.003	0.036
Late join	CC + MLO	0.788 ± 0.008	
Cross-view transformer (tokens)	CC + MLO	0.803 ± 0.007	0.061
Cross-view transformer (pixels)	CC + MLO	0.801 ± 0.003	0.006

Table 1: Area under the ROC curve for the CBIS-DDSM dataset. Mean and standard deviation computed over three runs. p-values for a two-sided Wilcoxon signed-rank test against the late-join baseline model.

Baseline: Late join network

Merge information from views only after global pooling.



Additional experiments

CheXpert chest X-ray with frontal and lateral views.

Conclusion

Combining multi-view information on a spatial level can outperform merging features after global pooling.

This research is part of the MARBLE project, funded by EFRO/OP-Oost.
MICCAI 2021, September 2021

Data Science Group, Faculty of Science, Radboud University, Nijmegen, the Netherlands
www.ru.nl/das · g.vantulder@cs.ru.nl

Radboud University
Nijmegen, the Netherlands