Attention-Gated Networks for Improving Ultrasound Scan Plane Detection

Jo Schlemper, Ozan Oktay, Liang Chen, Jacqueline Matthew,
Caroline Knight, Bernhard Kainz, Ben Glocker, Daniel Rueckert

BioMedIA Imperial College London

Motivation

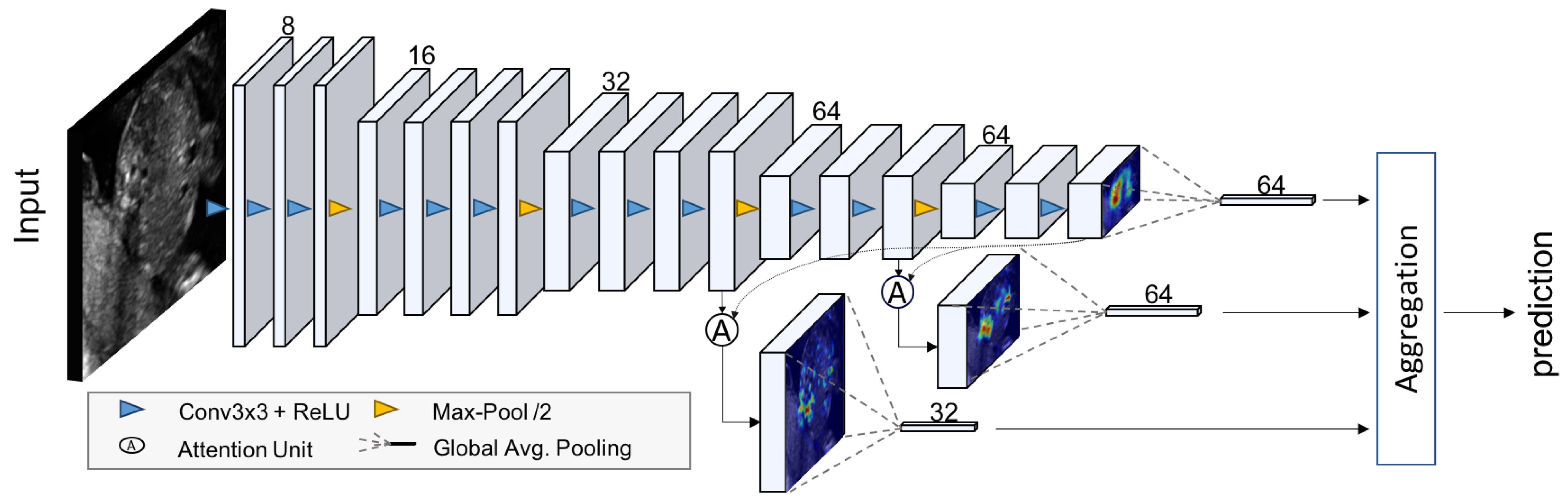
Scan plane detection in fetal ultrasound is a challenging problem due the poor image quality resulting in low interpretability for both clinicians and automated algorithms. In this work, we apply an *attention-gated network*[1] to *real-time* automated scan plane detection.

Contributions: We demonstrate:

- the superior classification performance of the network with a self-gated, soft-attention mechanism over the baseline,
- the proposed attention unit provides fine-scale attention maps that can be visualised with minimal computational overhead, and
- the attention maps can used for fast (weakly-supervised) object localisation.

Architecture

We use *Sononet*[2] as the backbone architecture and incorporate attention units at layer 11 and 14:



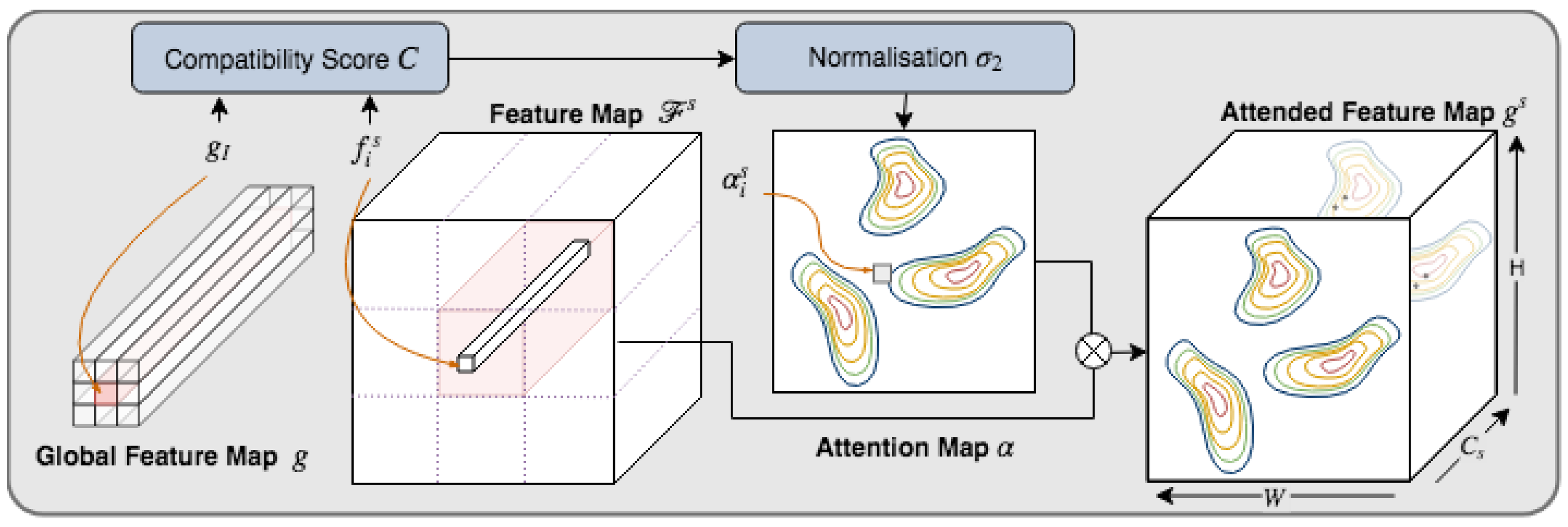
Each attended feature map with C channels is reduced to a vector $\mathbf{g}^s \in \mathbb{R}^{C_s}$ via global average pooling. **Aggregation strategy:** the network is first trained to predict at each attended scale via deep supervision. Secondly, an FC layer is fitted to the concatenated vector for futher finetuning.

Attention Unit

Let $\mathbf{f}_i^s$ represents the feature vector of pixel i at layer s, $\mathbf{g}$ be a global feature vector extracted just before the final soft-max layer. A *compatibility score* $\{c_i^s\}_{i=1}^n$ is defined as:

$$c_i^s = \mathbf{\Psi}\sigma_1\left(\mathbf{W}_f\mathbf{f}_i^s + \mathbf{W}_g\mathbf{g} + \mathbf{b}_g\right) + \mathbf{b}_\psi, \tag{1}$$

where $\{\mathbf{\Psi}, \mathbf{W}_f, \mathbf{W}_g, \mathbf{b}_g, \mathbf{b}_\psi\}$ are learnable and $\sigma_1(x) = \max(0, x)$. Once c_i^s's are computed, they are normalised across the feature map as $\alpha_i^s = (c_i^s - c_{\text{min}}^s)/\sum_i(c_i^s - c_{\text{min}}^s)$. Finally, a weighted sum $\mathbf{g}^s = \sum_{i=1}^n \alpha_i^s\mathbf{f}_i^s$ is computed, which aggregates the *attended* regions of the feature map.



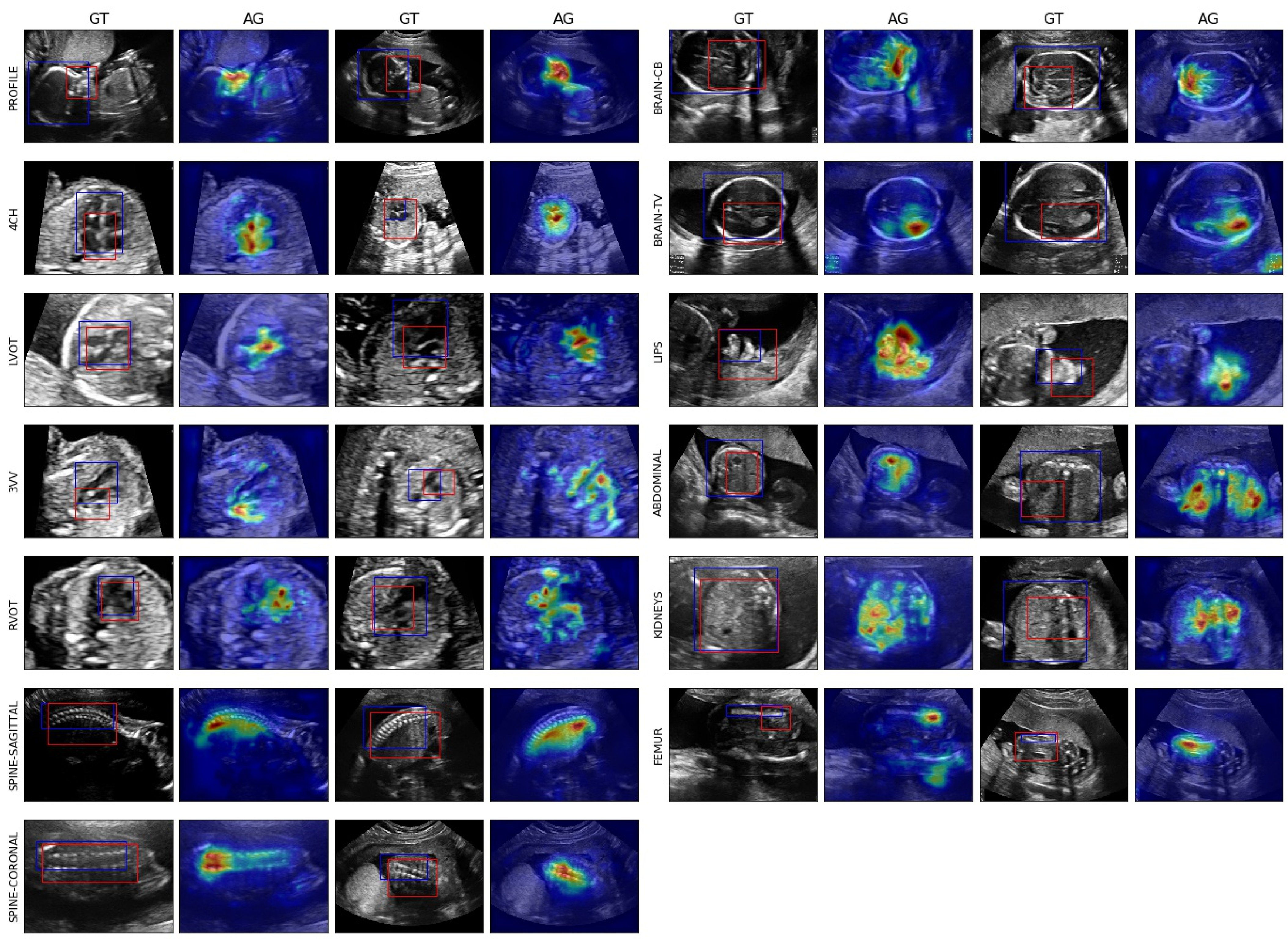
Results

Dataset: 2694 2D ultrasound examinations of volunteers with gestational ages between 18 and 22 weeks, containing 13 standard scan planes. The data was split into 122,233 training, 30,553 validation and 38,243 testing images.

Classification: Sononet (baseline) vs. AG-Sononet (proposed). "-n" shows the network size.

Method	Precision	Recall	Fwd/Bwd (ms)	#parameters
Sononet-8	0.878	0.922	1.36/2.60	0.16M
AG-Sononet-8	**0.916**	**0.929**	1.92/3.47	0.18M
Sononet-16	0.916	0.931	1.45/3.92	0.65M
AG-Sononet-16	**0.924**	**0.934**	1.94/5.13	0.70M
Sononet-32	0.924	**0.938**	2.40/6.72	2.58M
AG-Sononet-32	**0.931**	0.935	2.92/8.68	2.79M

Attention maps and weakly supervised localisation: (Blue: GT, Red: AG-Sononet)



Class-wise Improvement

Sononet-8 vs. AG-Sononet-8, the percentage change is shown in bracket:

View	Precision	Recall
Brain (Cb.)	0.99	0.98
Brain (Tv.)	0.98	0.99
Profile	0.95 **(0.06)**	0.96
Lips	0.98 **(0.03)**	0.96
Abdominal	0.96 (0.01)	0.96 (0.01)
Kidneys	0.86 **(0.05)**	0.90
Femur	0.99 (0.02)	0.98 (-0.01)
Spine (Cor.)	0.94 **(0.05)**	0.98
Spine (Sag.)	0.94 **(0.06)**	0.98 (-0.01)
4CH	0.94 **(0.04)**	0.97 (0.01)
3VV	0.69 **(0.05)**	0.72 (-0.01)
RVOT	0.69 **(0.03)**	0.71 **(0.04)**
LVOT	0.93 **(0.02)**	0.93 **(0.03)**
Background	0.99	0.99 (0.01)

Conclusion

The attention-gated network noticably improved precision with minimal additional parameters. It also enabled generating fine-grained attention map, which is explainable and can be used for weakly supervised object localisation.

Future work: A principled approach to incorporating and aggregating the attention-units will be investigated for improving training stability and performance.

References

[1] S. J. et al., "Learn to pay attention," in *International Conference on Learning Representations*, 2018.

[2] B. C. et al., "Sononet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound," *IEEE TMI*, 2017.