

# Processing Megapixel Images with Deep Attention-Sampling Models

Angelos Katharopoulos & François Fleuret

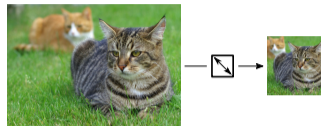
ICML, June 11, 2019



Funded by  FNSNF

# How do DNNs process large images?

Cropping and downsampling to a manageable resolution (e.g.  $224 \times 224$ )



Dividing the image into patches and processing them separately



\*image taken from the Imagenet dataset

## Our contributions

- ▶ **Sample from a soft attention** to only process a **fraction of the image** in high resolution.
- ▶ Derive **gradients through the sampling** for all parameters which allows to train our models end-to-end.
- ▶ Disentangle the computational and memory requirements from the input resolution.

# Soft Attention

Given an input  $x$  we define a neural network  $\Psi(x)$  that uses attention

$$\Psi(x) = g \left( \sum_{i=1}^K a(x)_i f(x)_i \right) = g \left( \mathbb{E}_{l \sim a(x)} [f(x)_l] \right),$$

where  $f(x) \in \mathbb{R}^{K \times D}$  are the features and  $a(x) \in \mathbb{R}_+^K$  is the attention distribution.

# Attention Sampling

We approximate  $\Psi(x)$  by Monte Carlo

$$\Psi(x) \approx g \left( \frac{1}{N} \sum_{q \in Q} f(x)_q \right) \text{ where } Q = \{q_i \sim a(x) \mid i \in \{1, 2, \dots, N\}\}.$$

We show that

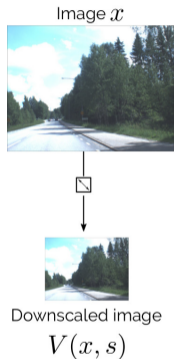
- ▶ Sampling from the attention is optimal to approximate  $\Psi(x)$  if  $\|f(x)_i\| = \|f(x)_j\| \forall i, j$
- ▶ We can compute the gradients both for the parameters  $a(\cdot)$  and  $f(\cdot)$

# Processing Megapixel Images with Deep Attention-Sampling Models

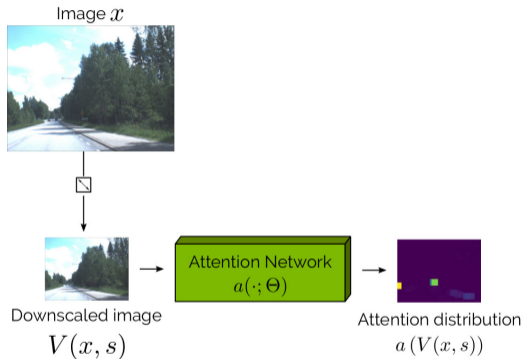
Image  $x$



# Processing Megapixel Images with Deep Attention-Sampling Models

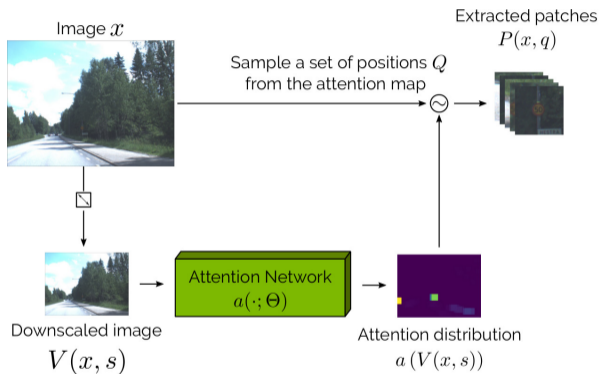


# Processing Megapixel Images with Deep Attention-Sampling Models

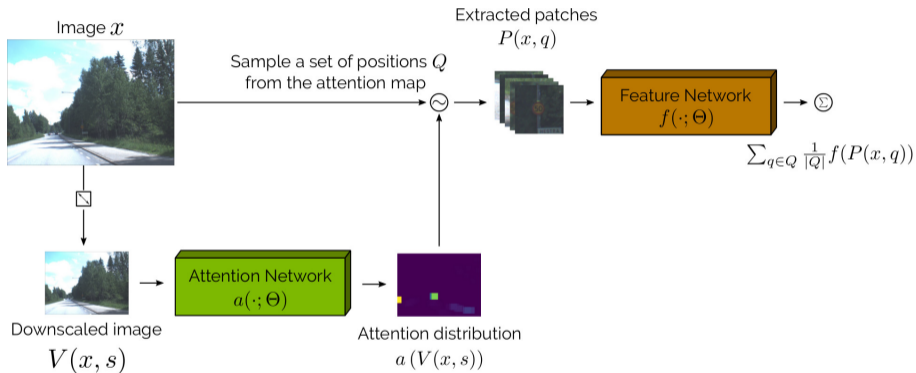




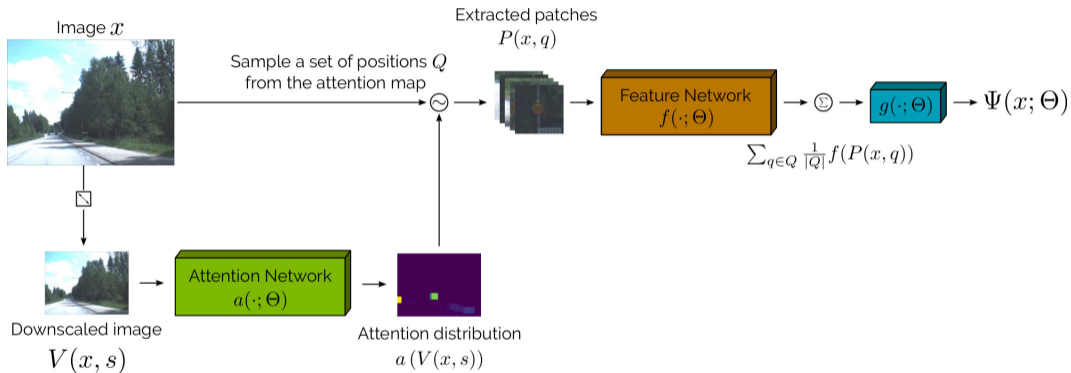
# Processing Megapixel Images with Deep Attention-Sampling Models



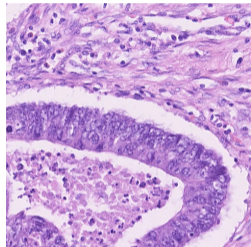
# Processing Megapixel Images with Deep Attention-Sampling Models



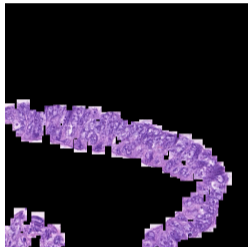
# Processing Megapixel Images with Deep Attention-Sampling Models



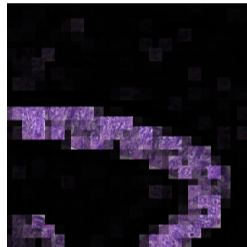
## Qualitative evaluation of the attention distribution (1)



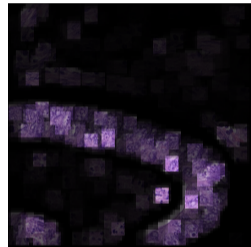
Full Image



Epithelial Cells



Ilse et al. (2018)



Attention Sampling

## Qualitative evaluation of the attention distribution (2)



Ground Truth



Ilse et al. (2018)



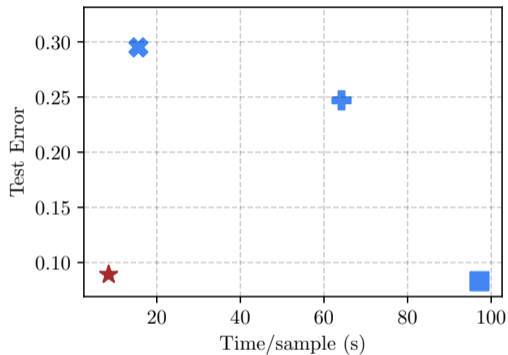
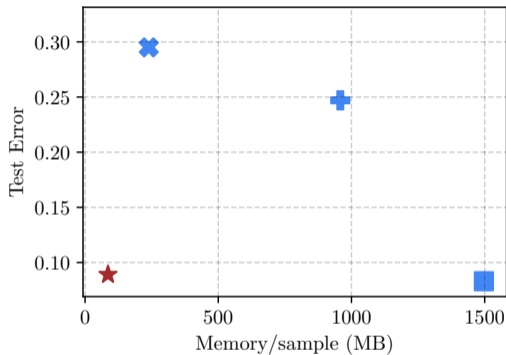
Attention Sampling



Extracted patch

# Thank you for your time!

## Speed limit sign detection



✱ CNN  $scale = 0.5$    ✚ CNN  $scale = 1.0$    ■ Ilse et al. 2018   ★ ATS (ours)

Come talk to us at **poster #3** at **Pacific Ballroom**.