

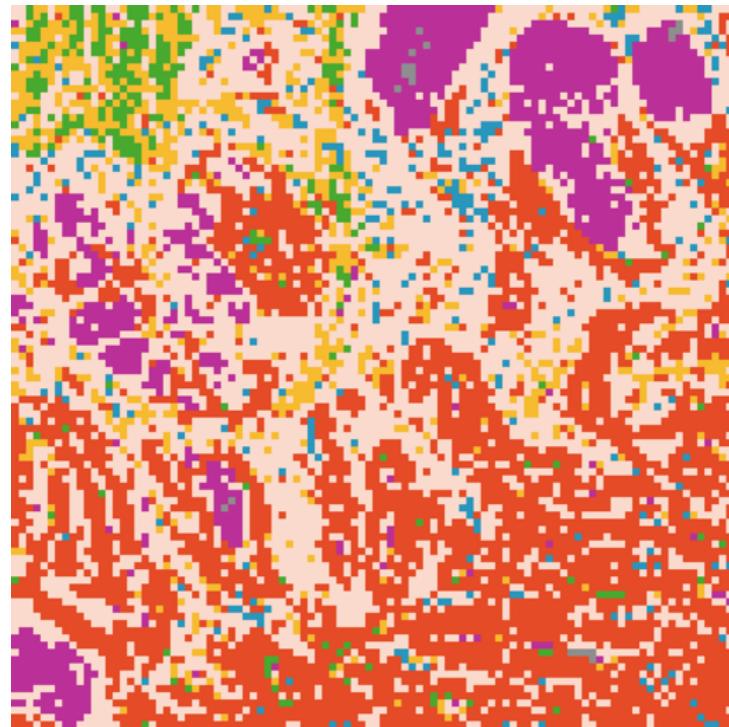
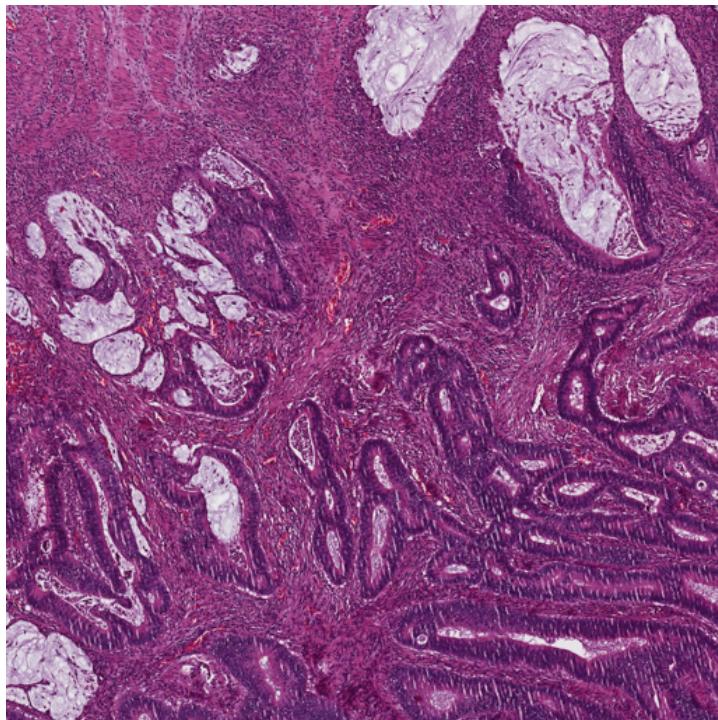
**ARA: accurate, reliable and active
histopathological image classification
framework with Bayesian deep learning**



Presentation outline

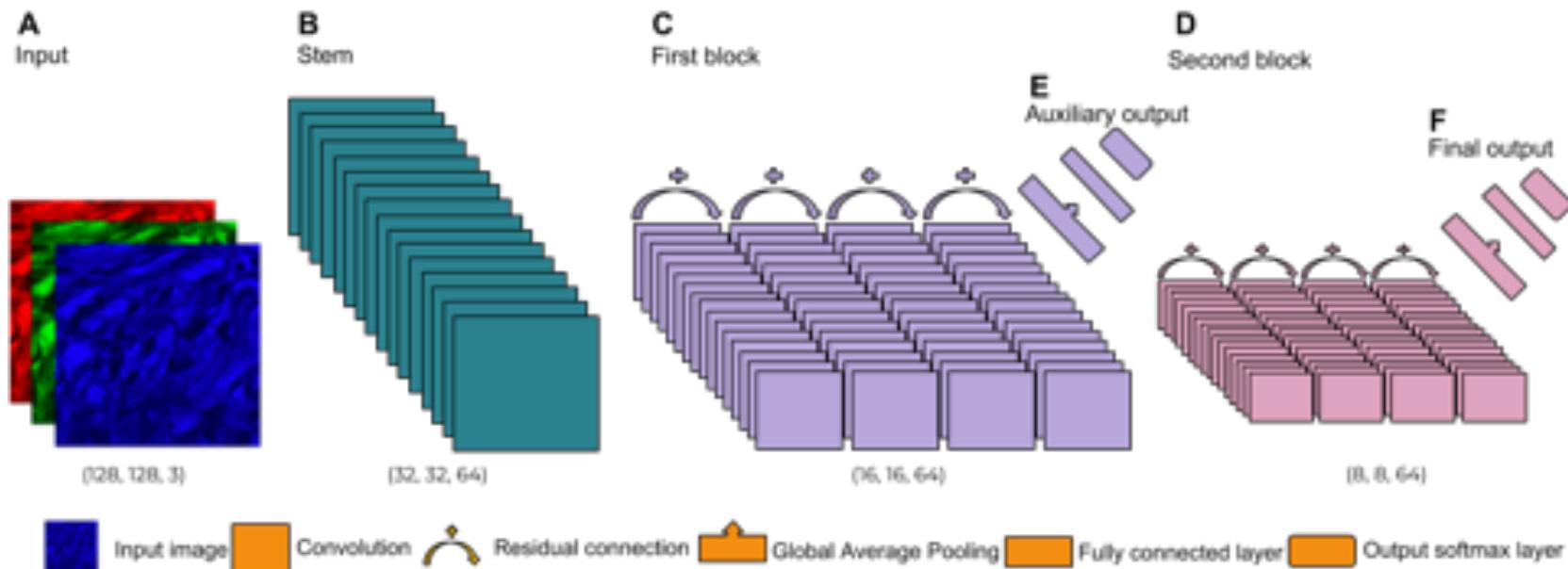
1. Project overview
2. ARA framework
 1. ResNets
 2. Dropout
 3. Variational dropout
 4. Framework overview
3. Kather et al. dataset results

Project overview

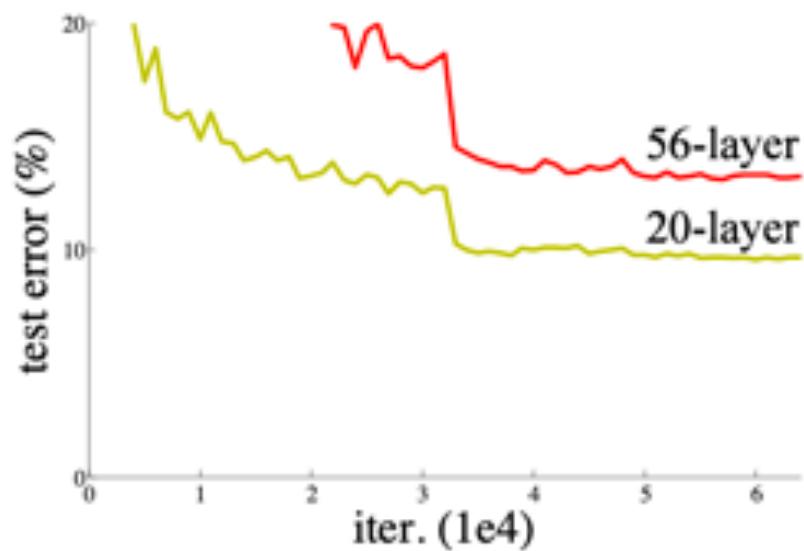
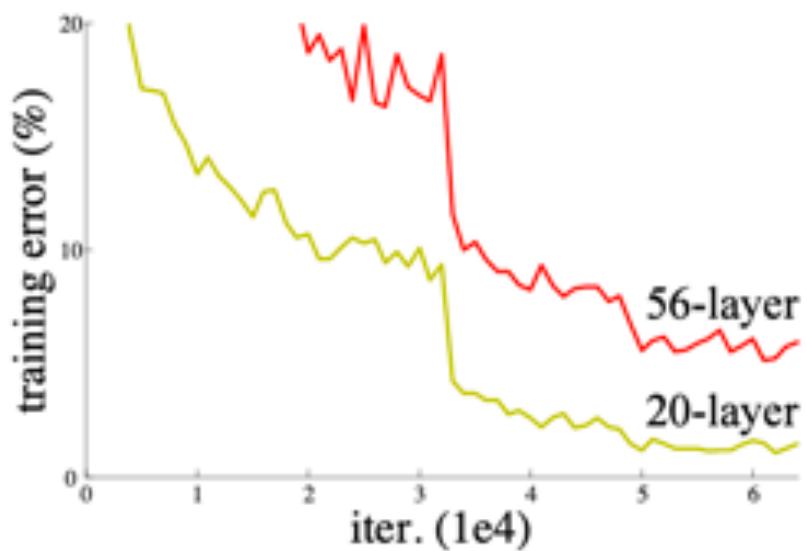


ARA framework

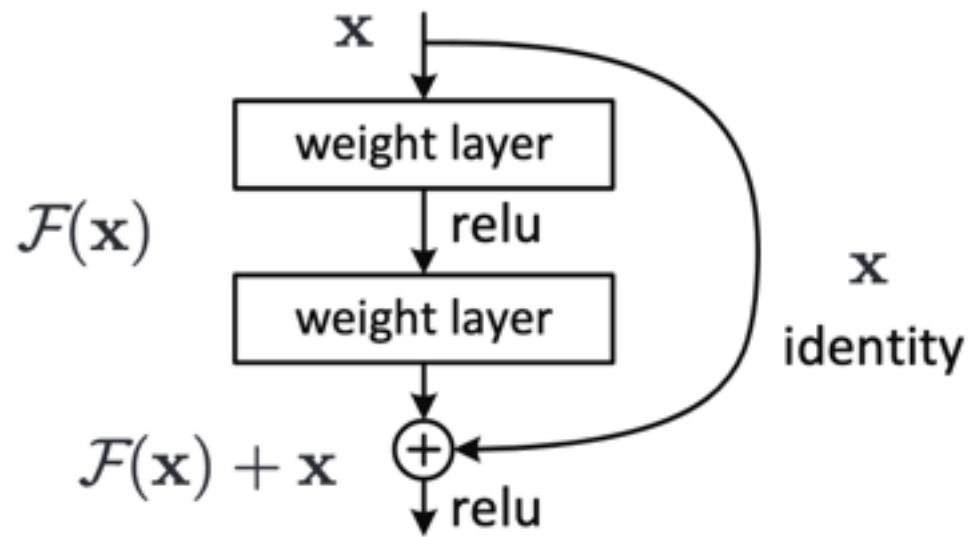
ARA-CNN architecture



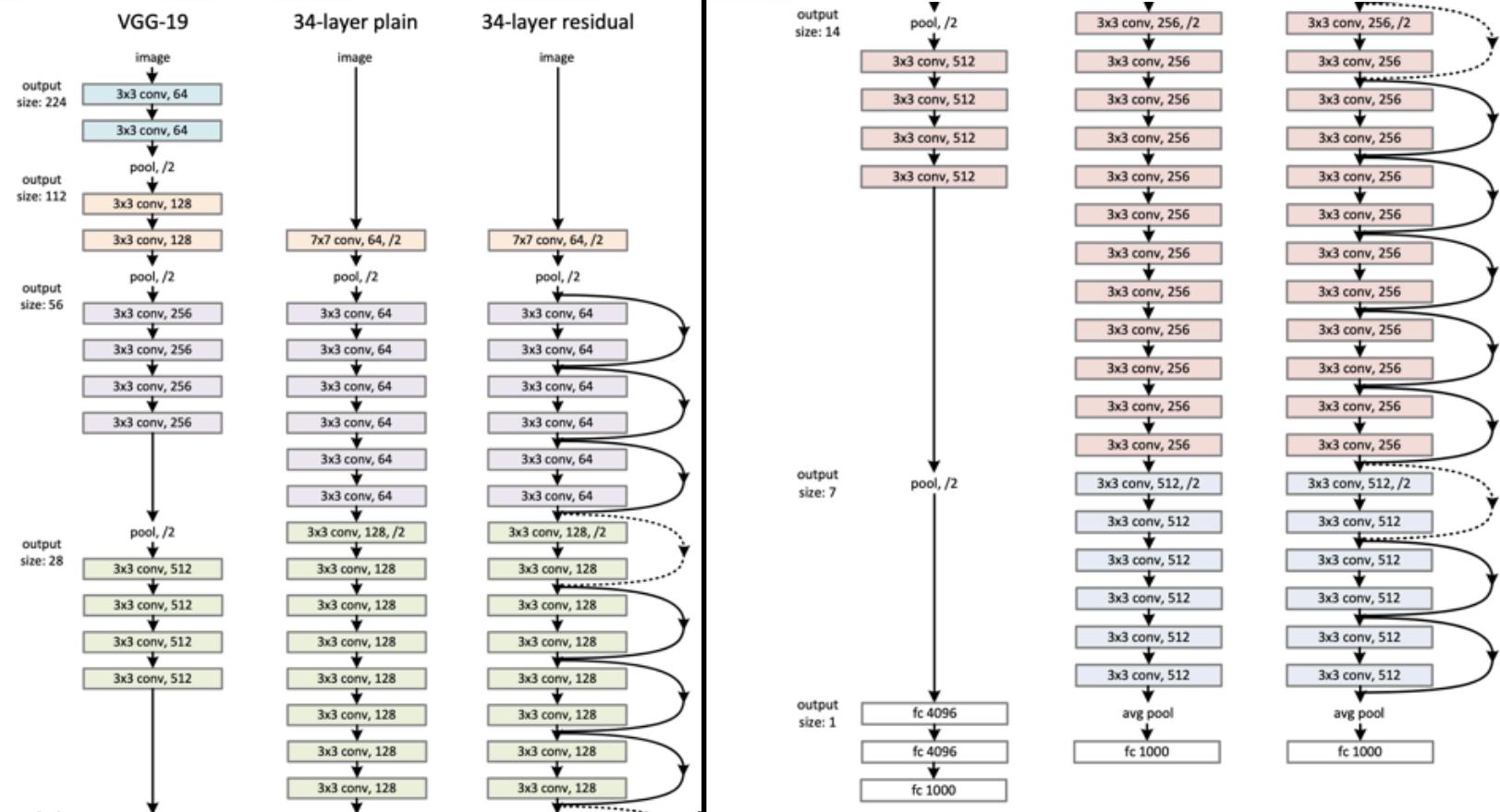
ResNet



ResNet

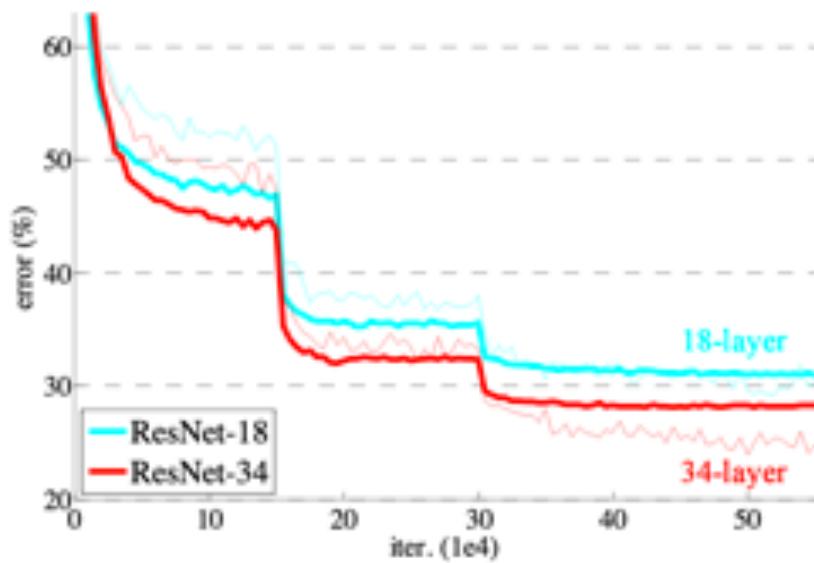
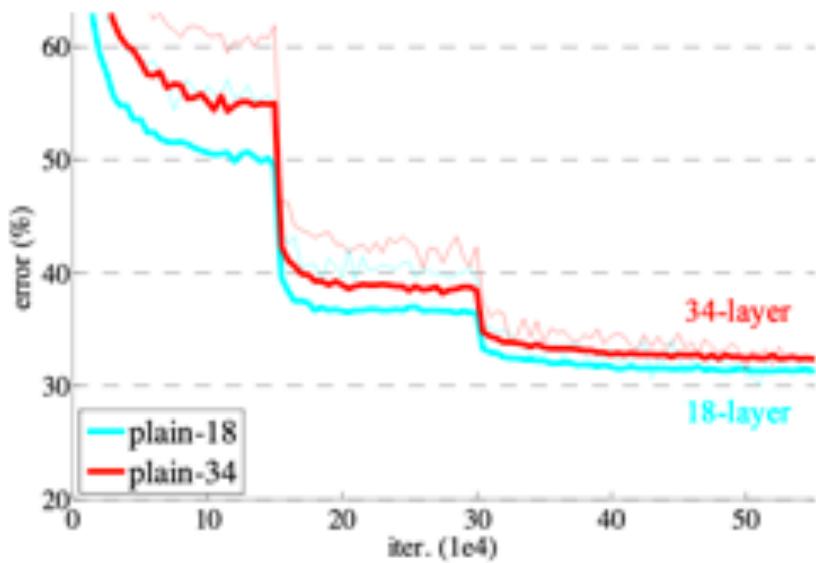


ResNet



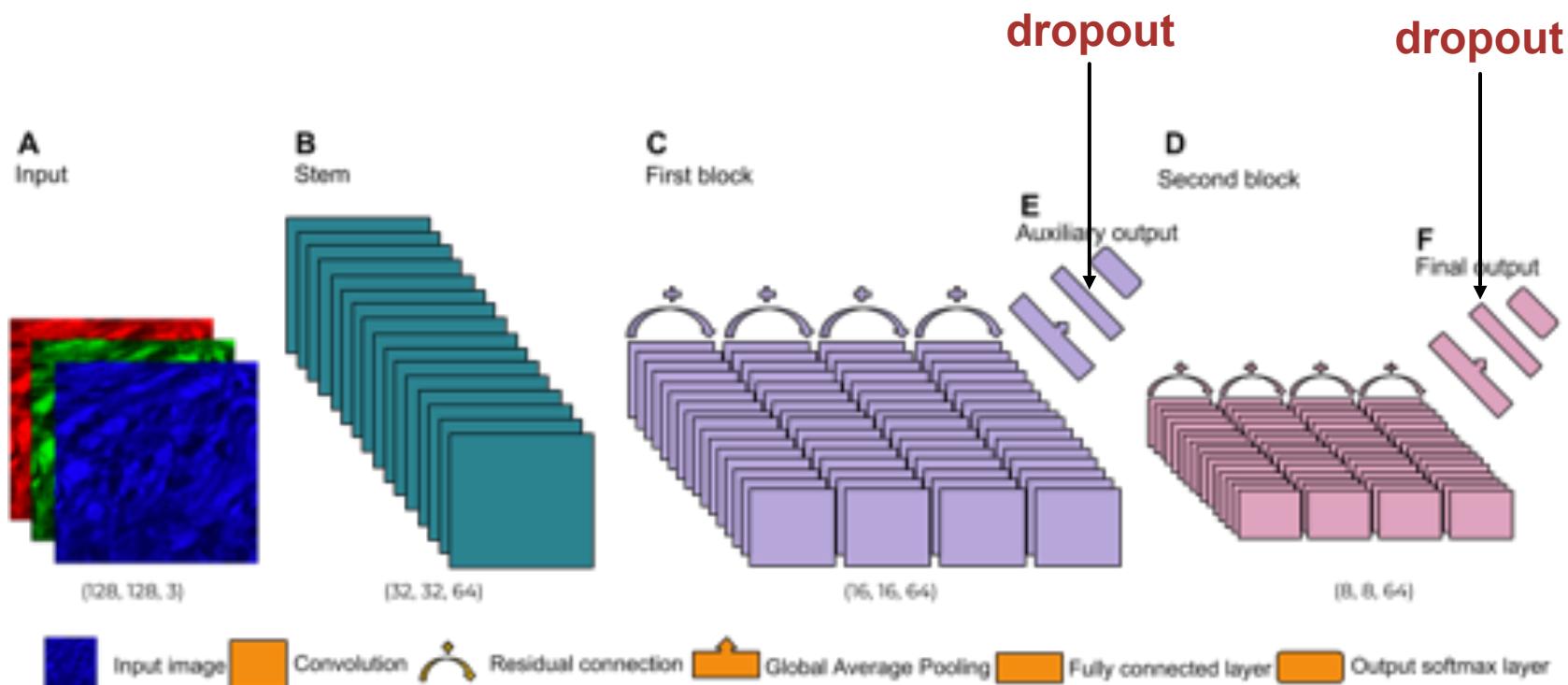
Deep Residual Learning for Image Recognition, He et al., 2015, arXiv:1512.03385

ResNet

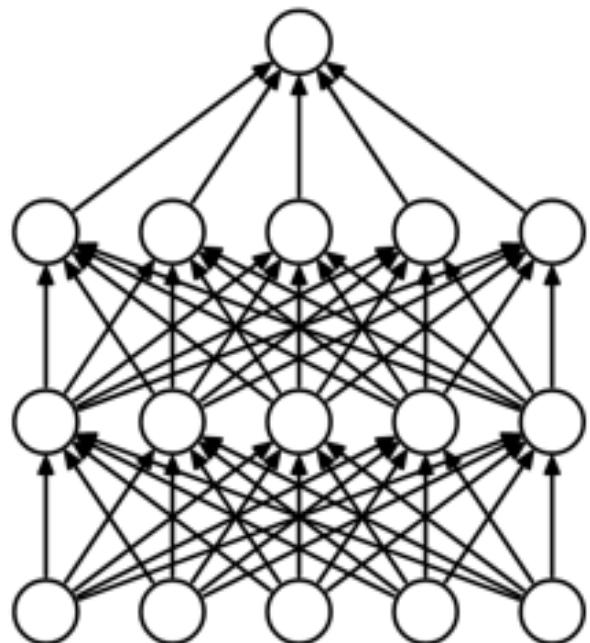


ARA framework

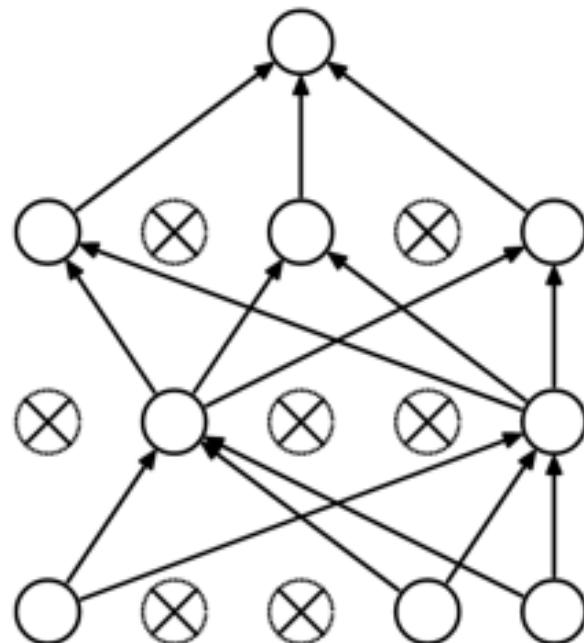
ARA-CNN architecture



Dropout



(a) Standard Neural Net



(b) After applying dropout.

Dropout: A simple way to prevent neural networks from overfitting, Srivastava et al., JMLR 15, 2014

Variational dropout

*Dropout as a Bayesian Approximation:
Representing Model Uncertainty in Deep Learning,
Gal, Ghahramani, 2015, arXiv:1506.02142*

- Notation:
 - Training data D , observation o , class $y^* \in \{1 \dots M\}$, M classes
 - Parameter matrix ω ,
 - $P(y^*|o, \omega)$ is the probability of assigning o to class y^*
 - $\hat{\omega}_t$ are parameters for the network from variational dropout call t , $t = 1, \dots, T$

Variational dropout

*Dropout as a Bayesian Approximation:
Representing Model Uncertainty in Deep Learning,
Gal, Ghahramani, 2015, arXiv:1506.02142*

- Notation:
 - Training data D , observation o , class $y^* \in \{1 \dots M\}$, M classes
 - Parameter matrix ω ,
 - $P(y^*|o, \omega)$ is the probability of assigning o to class y^*
 - $\hat{\omega}_t$ are parameters for the network from variational dropout call t , $t = 1, \dots, T$
- Dropout is active during inference
- We assume parameters ω are random variables and their distribution *a posteriori* $P(\omega | D)$ is approximated by q , so $\hat{\omega}_t \sim q(\omega)$
- *Predictive distribution:*

$$P(y^*|o, D) = \int P(y^*|o, \omega)P(\omega|D)d\omega \approx \int P(y^*|o, \omega)q(\omega)d\omega \approx \frac{1}{T} \sum_{t=1}^T P(y^*|o, \hat{\omega}_t).$$

Variational dropout

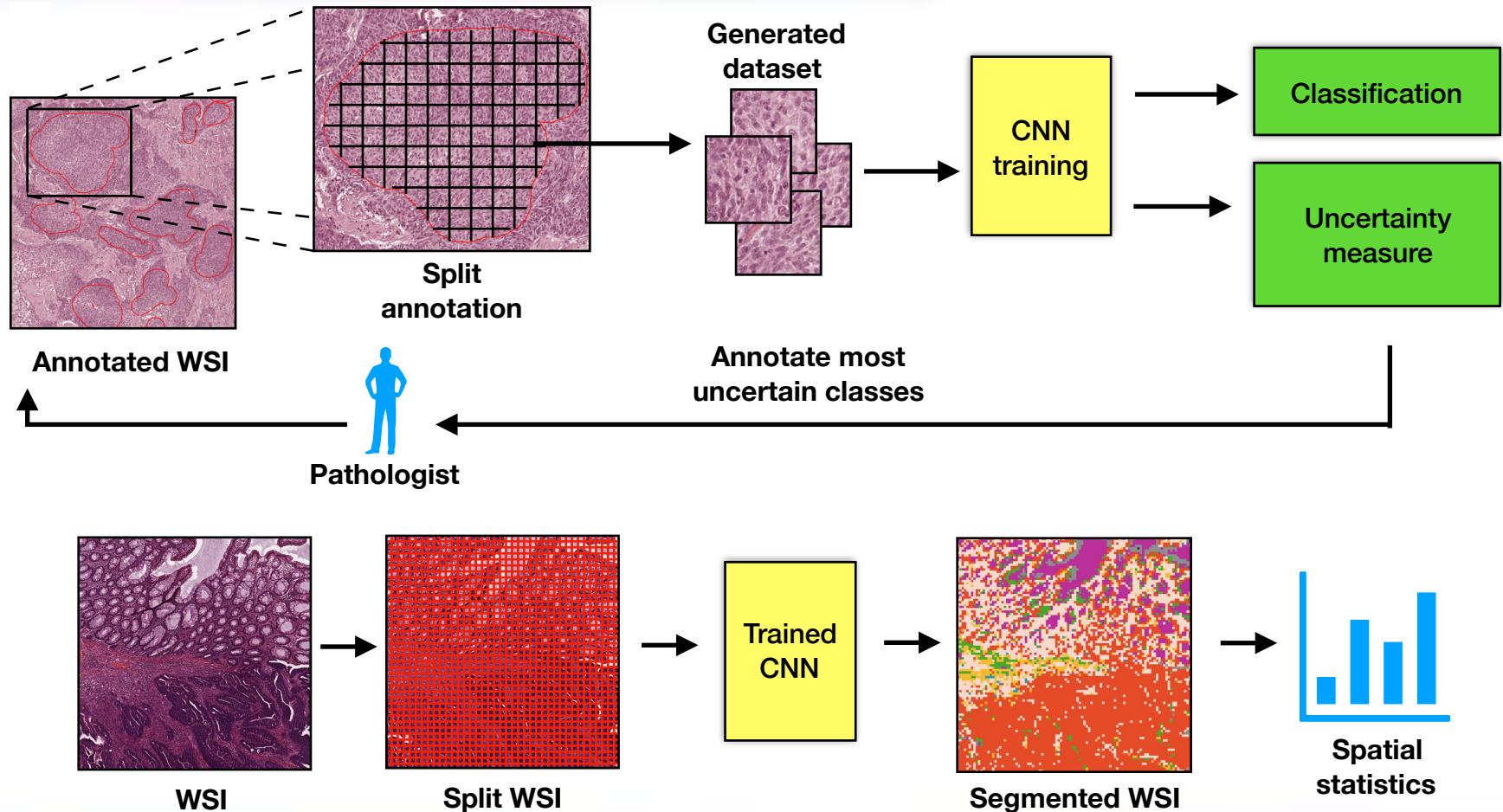
*Dropout as a Bayesian Approximation:
Representing Model Uncertainty in Deep Learning,
Gal, Ghahramani, 2015, arXiv:1506.02142*

- Notation:
 - Training data D , observation o , class $y^* \in \{1 \dots M\}$, M classes
 - Parameter matrix ω ,
 - $P(y^*|o, \omega)$ is the probability of assigning o to class y^*
 - $\hat{\omega}_t$ are parameters for the network from variational dropout call t , $t = 1, \dots, T$
- Dropout is active during inference
- We assume parameters ω are random variables and their distribution *a posteriori* $P(\omega | D)$ is approximated by q , so $\hat{\omega}_t \sim q(\omega)$
- *Predictive distribution:*
$$P(y^*|o, D) = \int P(y^*|o, \omega)P(\omega|D)d\omega \approx \int P(y^*|o, \omega)q(\omega)d\omega \approx \frac{1}{T} \sum_{t=1}^T P(y^*|o, \hat{\omega}_t).$$
- Uncertainty is the entropy of distribution $P(y^*|o, D)$

$$H[P(y^*|o, D)] = - \sum_{y^* \in \{1 \dots M\}} P(y^*|o, D) \log P(y^*|o, D)$$

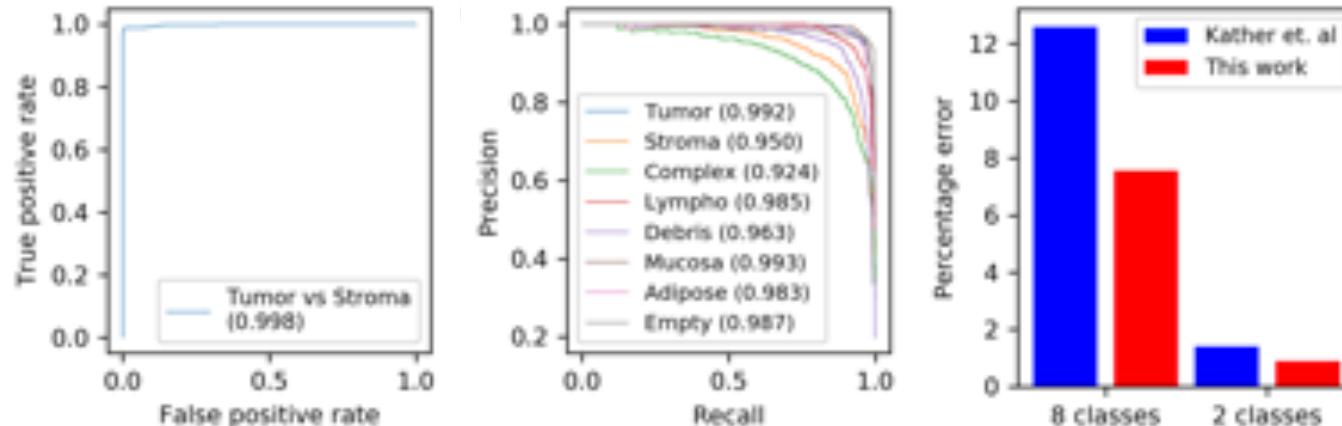
ARA framework

ARA overview



Kather et al. dataset results

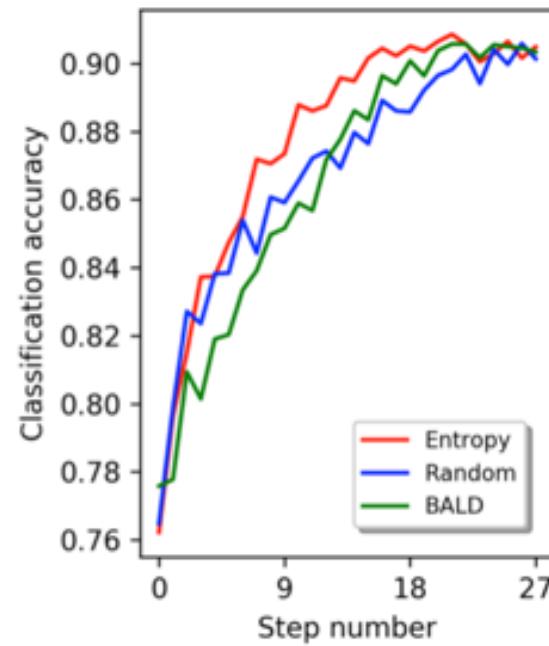
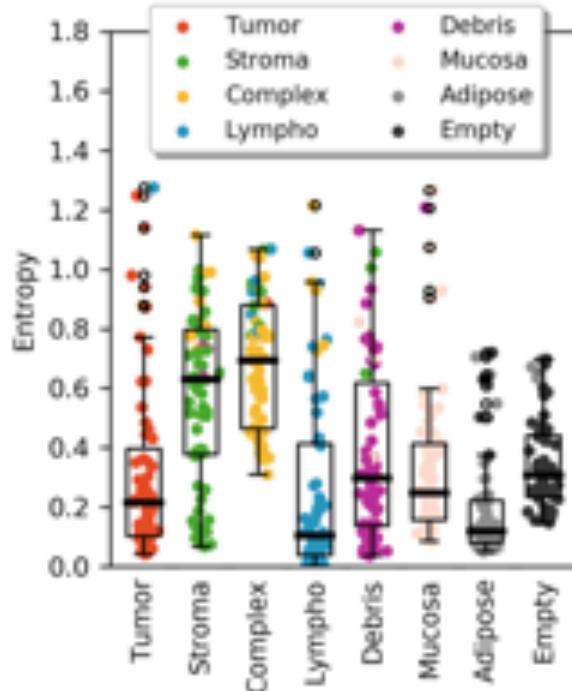
Classification accuracy



| Method | Method type | Problem type | Max. reported 10-fold ACC | Max. reported 5-fold ACC | Max. reported 2-fold ACC | 10-fold AUC | 5-fold AUC |
|-----------------|-------------|--------------|--------------------------------------|--------------------------------------|-----------------------------|----------------|---------------|
| Kather et al. | Traditional | Binary | 98.6% | - | - | - | - |
| | | Multiclass | 87.4% | - | - | 0.976 | - |
| Ribeiro et al.* | Traditional | Binary | 97.68% | - | - | - | - |
| Sarkar et al. | Traditional | Multiclass | 73.66% | - | - | - | - |
| Wang et al. | CNN | Multiclass | - | $92.6 \pm 1.2\%$ | - | - | 0.985 |
| Pham | CNN | Binary | - | - | 84.00% | - | - |
| ARA-CNN | CNN | Binary | $99.11 \pm 0.97\%$ | $98.88 \pm 0.52\%$ | 98.88% | 0.998 | 0.999 |
| | | Multiclass | $92.44 \pm 0.81\%$ | $92.24 \pm 0.82\%$ | $88.92 \pm 1.95\%$ | 0.995 | 0.995 |

Kather et al. dataset results

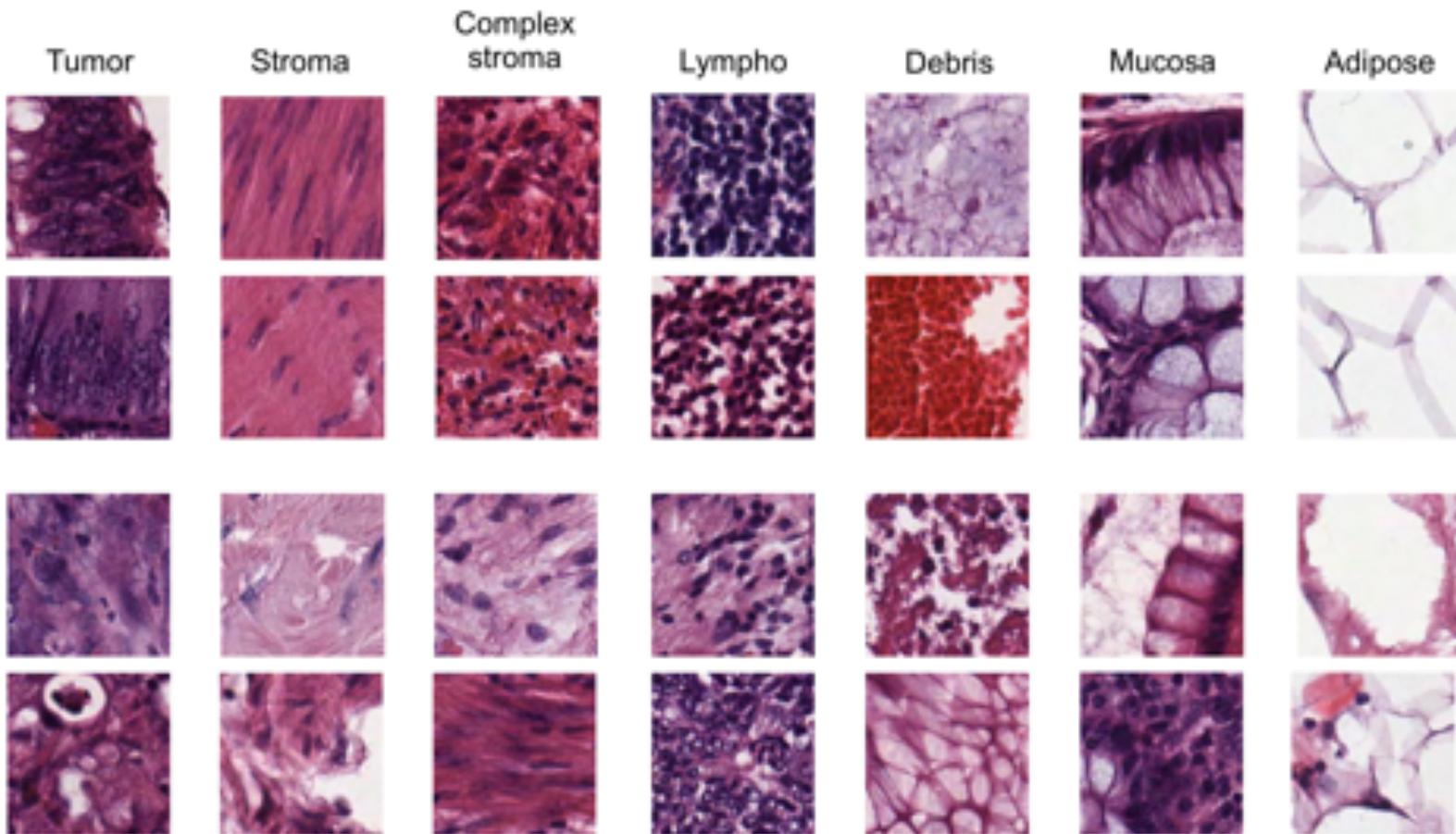
Uncertainty analysis and active learning



Kather et al. dataset results

Uncertainty analysis

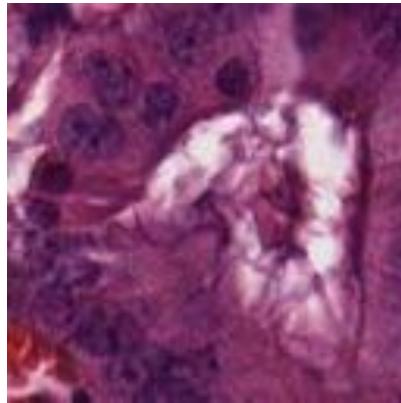
Low uncertainty H



High uncertainty H

Kather et al. dataset results

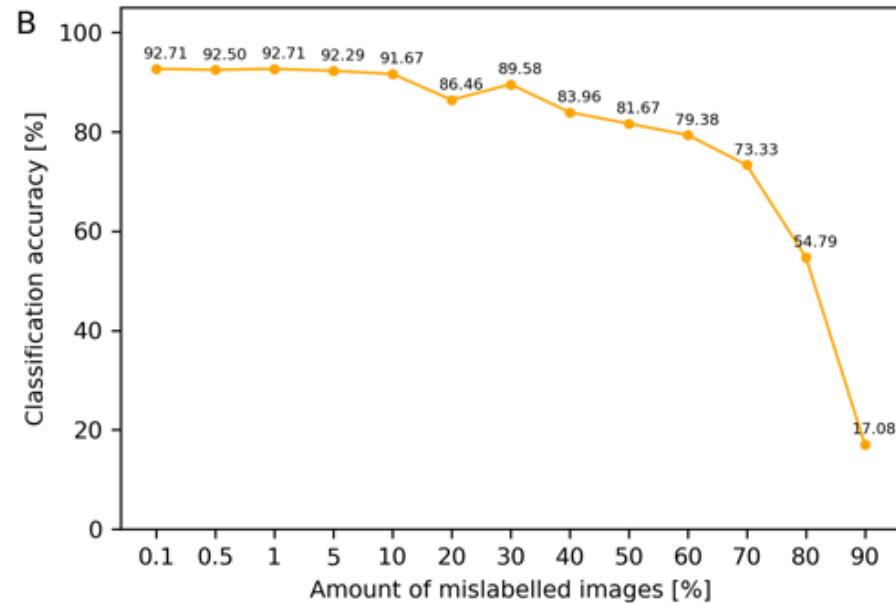
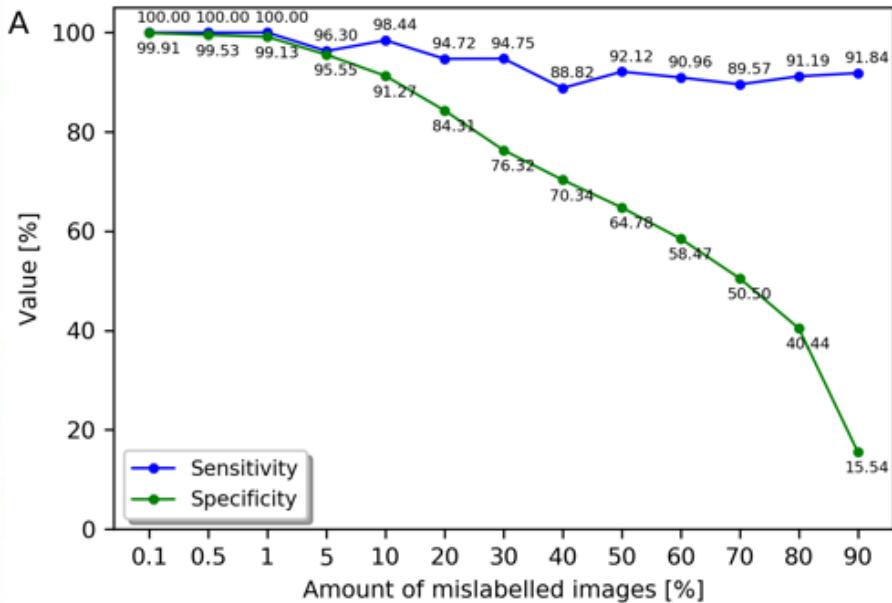
Identification of mislabelled samples



| Label in dataset | Actual label |
|------------------|--------------|
| Stroma | Tumour |

Kather et al. dataset results

Identification of mislabelled samples



Kather et al. dataset results

Paper

ARA: accurate, reliable and active histopathological image classification framework with Bayesian deep learning

Łukasz Rączkowski¹, Marcin Możejko¹, Joanna Zambonelli², and Ewa Szczęska^{1,*}

¹Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Warsaw, Poland

²Department of Pathology, Medical University of Warsaw, Warsaw, Poland

Scientific Reports 9, Article number: 14347 (2019), doi:10.1038/s41598-019-50587-1

Thank you for your attention!

