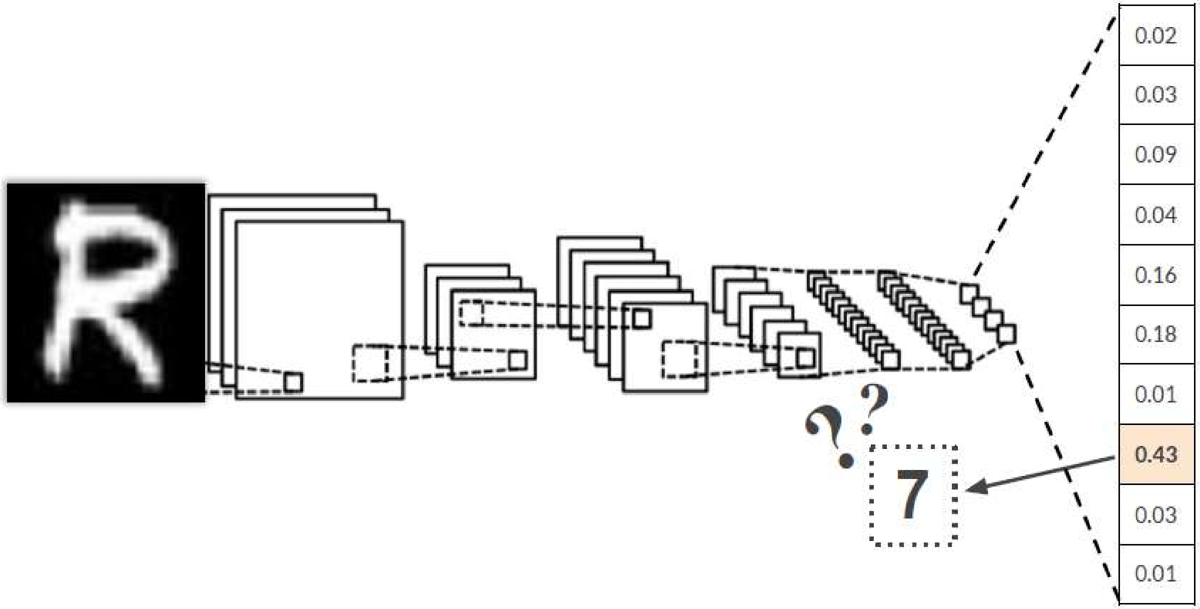
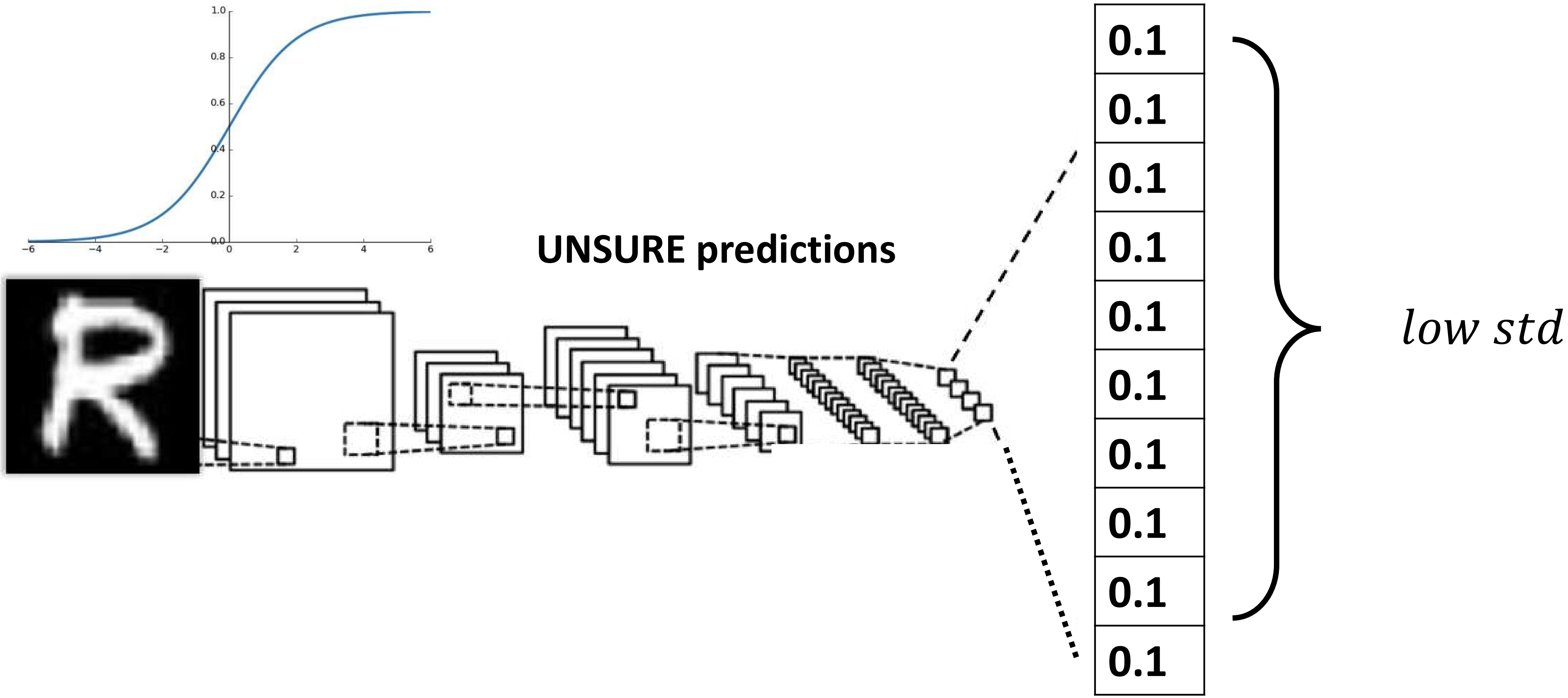
POLITECNICO DI TORINO

How to make your CNN say I don't know?



POLITECNICO DI TORINO

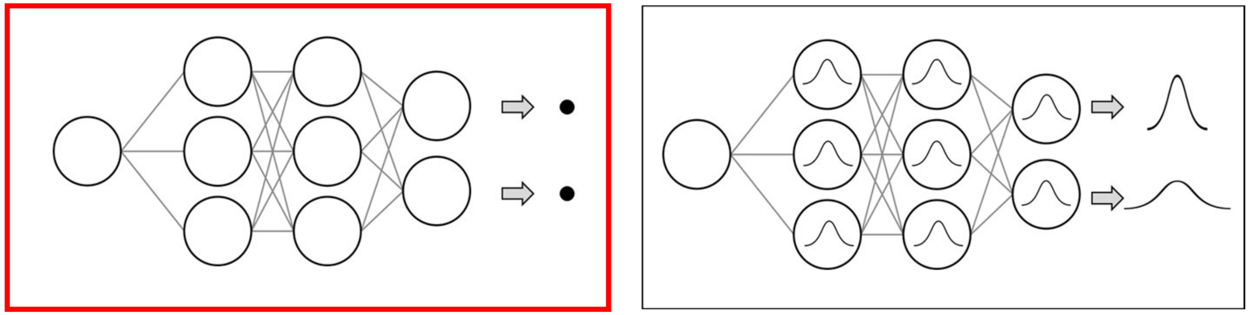
Softmax ideal behaviour



POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?

- Bayesian modelling: the **weights' distribution**.

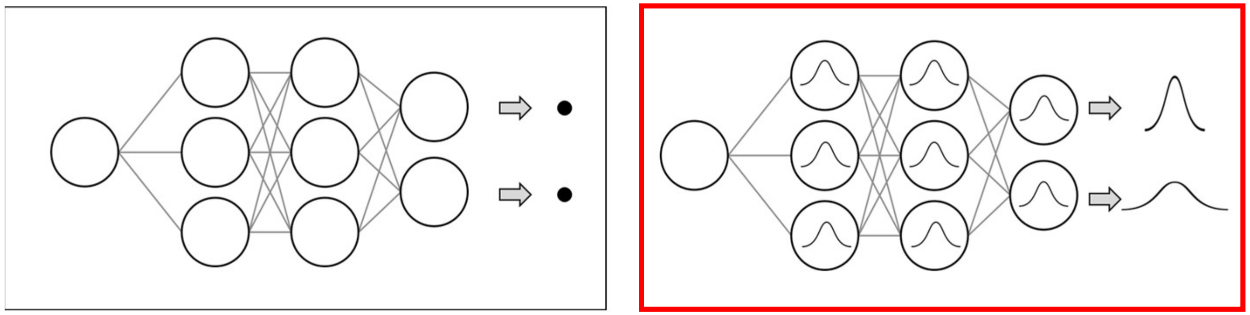


Standard CNNs: weights are POINT ESTIMATE VALUES

POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?

- Bayesian modelling: the **weights’** distribution.



B-CNNs: weights described by PDFs computed as BAYESIAN POSTERIORS [1]

[1] Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning (Gal et al., 2016)

POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?

- We use Bayesian inference to estimate weights' PDFs
- Given a prior over the weights W of the CNN, the objective of Bayesian inference is to find a posterior distribution over all model's parameters W

$$p(\mathbf{W}|X,Y) = \frac{p(Y|X,\mathbf{W})p(\mathbf{W})}{p(Y|X)}$$

POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?

- We use Bayesian inference to estimate weights’ PDFs
- Given a prior over the weights W of the CNN, the objective of Bayesian inference is to find a posterior distribution over all model's parameters W

$$p(\mathbf{W}|X,Y) = \frac{p(Y|X,\mathbf{W})\,p(\mathbf{W})}{p(Y|X)}$$

Likelihood : how likely is the model with W given the input X

POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?

- We use Bayesian inference to estimate weights' PDFs
- Given a prior over the weights W of the CNN, the objective of Bayesian inference is to find a posterior distribution over all model's parameters W

$$p(\mathbf{W}|X, Y) = \frac{p(Y|X, \mathbf{W})p(\mathbf{W})}{p(Y|X)}$$

Prior: prior beliefs we have on the parameters

POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?

- We use Bayesian inference to estimate weights’ PDFs
- Given a prior over the weights W of the CNN, the objective of Bayesian inference is to find a posterior distribution over all model's parameters W

$$p(\mathbf{W}|X,Y) = \frac{p(Y|X,\mathbf{W})p(\mathbf{W})}{p(Y|X)}$$

$$p(Y|X) = \int p(Y|X,W)dw$$

Marginal likelihood integral: normalization constant to ensure that posterior respects PDF properties

POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?

- We use Bayesian inference to estimate weights' PDFs
- Given a prior over the weights W of the CNN, the objective of Bayesian inference is to find a posterior distribution over all model's parameters W

$$p(\mathbf{W}|X,Y) = \frac{p(Y|X,\mathbf{W})p(\mathbf{W})}{p(Y|X)}$$

$$p(Y|X) = \int p(Y|X,W)dw$$

Marginal likelihood integral: normalization constant to ensure that posterior respects PDF properties

Intractable for CNNs

Why Bayesian Convolutional Neural Networks (B-CNNs)?

- We need a way to approximate the real Bayesian posterior for weights W

- Variational inference

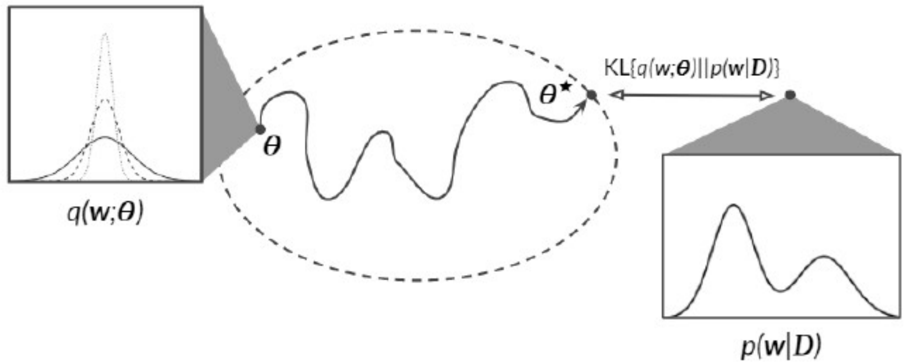
    i. We minimize the KL divergency between the posterior and a generic distribution easier which is to work with
    ii. Minimizing KL divergence is known to be equivalent to maximizing the so-called evidence lower bound (ELBO)

- <u>Variational inference turns the integration problem into an optimization one</u>

$$KL\{q_\theta(\omega)||p(\omega|X,Y)\} = \int_\Omega q_\theta(\omega) \log \frac{q_\theta(\omega)}{p(\omega|X,Y)} d\omega \quad \text{(i)}$$

$$\int_\Omega q_\theta(\omega) \log p(y|x,\omega) d\omega - KL\{q_\theta(\omega)||p(\omega)\} \quad \text{(ii)}$$

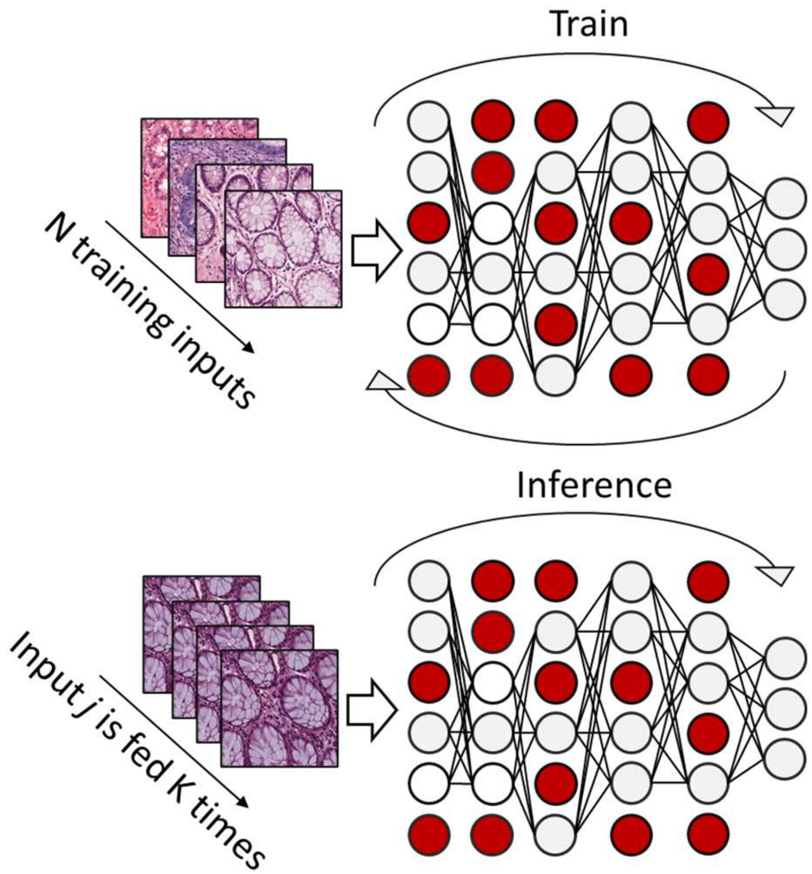
Ideally is 0



POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?

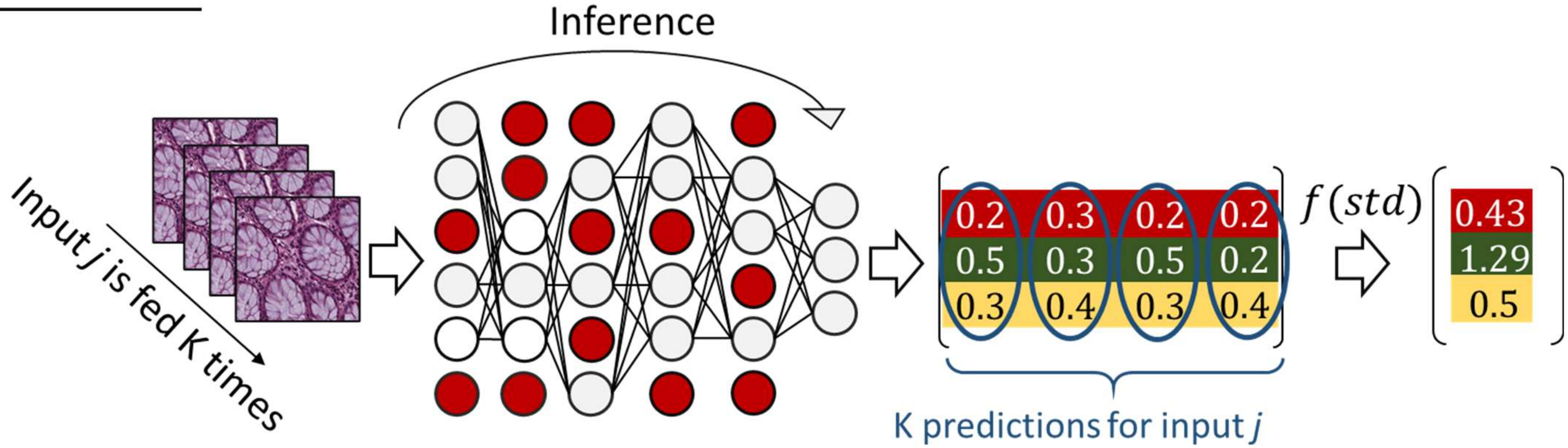
- We need a way to approximate the real Bayesian posterior for weights W
- Variational dropout method
- Consists in applying dropout before each trainable layer in a deep network, also at inference time
- This has been shown to be equivalent to gaussian distributions for the weights [1]



[1] Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning (Gal et al., 2016)

POLITECNICO DI TORINO

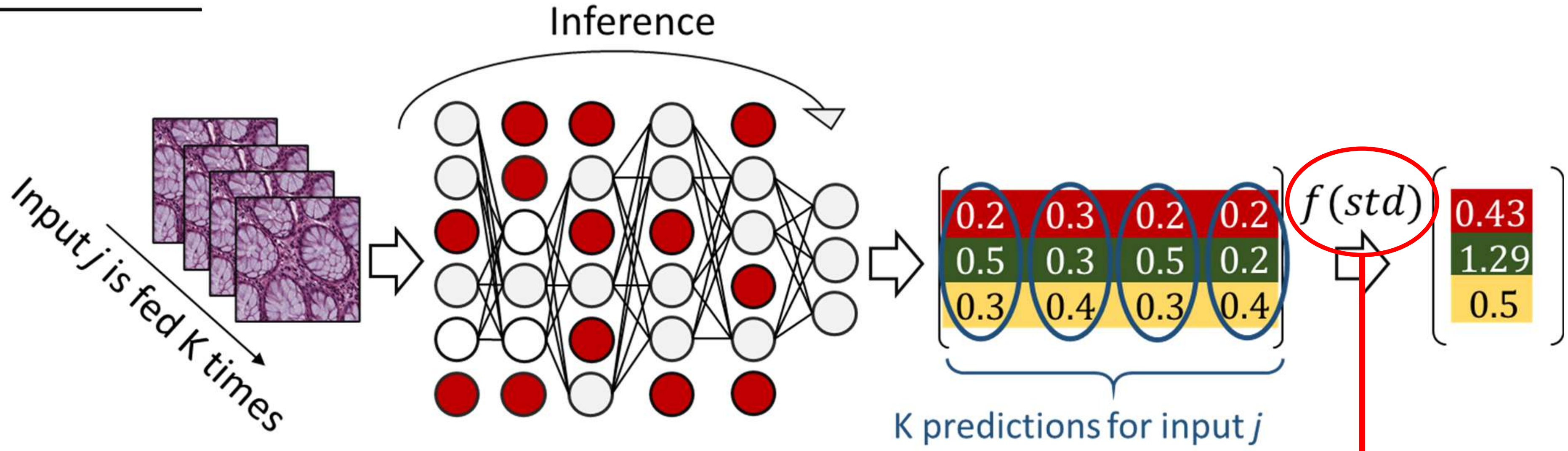
Why Bayesian Convolutional Neural Networks (B-CNNs)?



- At inference time, for the same input *j* we get K different outputs because the model uses different parameters every time
- From the prediction's distribution, an uncertainty measure based on standard deviation can be retrieved

POLITECNICO DI TORINO

Why Bayesian Convolutional Neural Networks (B-CNNs)?



- Uncertainty formula [2]:

$$\frac{1}{T}\sum_{t=1}^{T} diag(\hat{p}_t) - \hat{p}_t^{\otimes 2} + \frac{1}{T}\sum_{t=1}^{T}(\hat{p}_t - \bar{p})^{\otimes 2}$$

[2] Uncertainty quantification using Bayesian neural networks in classification: Application to biomedical image segmentation (Kwon et al., 2019)