STATISTICAL MASKING OF A VAE SEGMENTATION

Andres Fernandez
Department of Computer Science
Goethe University
Frankfurt am Main
`aferro@em.uni-frankfurt.de`

September 21, 2019

ABSTRACT

Given an input image $A \in \mathbb{R}^{H \times W}$, we have a variational classifier that maps every pixel $A_{i,j}$ to a normal distribution (parametrized by $\begin{pmatrix} \mu_{i,j} \\ \sigma_{i,j} \end{pmatrix}$) representing the confidence of the pixel belonging to a class. The output map has then dimensionality $B \in \mathbb{R}^{H \times W \times 2}$.

We want to accept or reject a given pixel based on this output. One method to do this is to set a minimum threshold on μ and a maximum threshold on σ, thus accepting predictions that are high and confident enough. This has several drawbacks:

- We have to set (and eventually look for) 2 thresholds.
- A normal sample with $\begin{pmatrix} \mu + \epsilon; \\ \sigma - \epsilon \end{pmatrix}$ (for arbitrarily small ϵ) will get accepted, but half of its total probability will be outside of our accepted range. This doesn't make the most of the information provided by the variational classifier.

A better criterium is to accept a prediction if **a probability of at least** $T \in (0, 1)$ **is contained above a threshold** $x \in \mathbb{R}$. This has the advantage that T can be fixed in beforehand because of its clear semantics, leaving the mask as a function of one single threshold only.

We show here that to fullfil that criterium it suffices to compute, for every pixel:

$$\mathbb{1}_{accept} = \frac{\mu - x}{\sigma} \geq prob(T) \tag{0.1}$$

Where $x \in \mathbb{R}$ is our threshold, and $\mathbb{1}_{accept} \in \{0, 1\}$ the boolean mask value.

1 Definitions

A normal distribution can be standarized as follows:

$$
\begin{aligned}
X &\sim \mathcal{N}(\mu, \sigma^2) \\
Y = \frac{X - \mu}{\sigma} &\sim \mathcal{N}(0, 1)
\end{aligned}
\tag{1.1}
$$

Following this, the cumulative distribution function (CDF) can also be standarized as follows:

$$CDF_{\mu,\sigma}(x) := P(X \leq x) = P(\sigma Y + \mu \leq x) = P(Y \leq \frac{x-\mu}{\sigma}) = CDF_{0,1}(\frac{x-\mu}{\sigma})$$
$$=: \Phi(\frac{x-\mu}{\sigma}) = t \tag{1.2}$$

By the same reasoning, the inverse of the CDF (called *quantile* function) can also be standarized as follows:

$$quant_{\mu,\sigma}(t) := CDF_{\mu,\sigma}^{-1}(t) = x$$
$$prob(t) := CDF_{0,1}^{-1}(t) = \frac{x-\mu}{\sigma} \iff quant_{\mu,\sigma}(t) = \sigma \cdot prob(t) + \mu \tag{1.3}$$

The quantile for the standard normal is called *probit*:

$$prob(t) = CDF_{0,1}^{-1}(t) = \Phi^{-1}(t) \tag{1.4}$$

2   Proof

Note how the CDF definitions in 1.2 provide semantics for the parameters (t, x): Given a distribution $\mathcal{N}(\mu, \sigma^2)$, the set of all events below threshold x will have a probability of t.

This corresponds closely to our specifications: we want to accept a distribution if a probability of at least $T \in (0, 1)$ is contained above a threshold x. Therefore criterium 0.1 can be easily derived as follows:

$$CDF_{\mu,\sigma}(x) \leq 1 - T \iff x \leq quant_{\mu,\sigma}(1-T) \iff x \leq \sigma \cdot prob(1-T) + \mu$$
$$\iff \frac{x-\mu}{\sigma} \leq prob(1-T) \tag{2.1}$$
$$\iff \frac{\mu-x}{\sigma} \geq prob(T) \qquad \square$$

Assuming there is an efficient way to compute $prob(T)$ (many programming libraries include one), this expression can be computed very efficiently. It is also fully vectorizable so it also scales up well for the whole image map with adequate hardware.