Unsupervised Lung Cancer nodule Representation using Region of Interest

March 13, 2020

1 Training objective

Let I be the image of dimension $n * n$. For now let the scale be fixed to $m * m$. The task is to train a prediction function g.

Let each block be indexed $x_1, x_2, x_3 ... x_n$.

$$f_\theta(x_i) + e_i \rightarrow v_i$$

where,
$x_i \in m \times m$
$v_i \in d$
e_i : positional embedding
f_θ : convnet with params θ
θ is trainable parameters.

Each block x_i is passed through a function f to get vector representation of each block, where $v_1, v_2, v_3 .... v_n$ represent vector representations.

Now, we apply a pooling layer p over all these vector representations, except the block to be predicted.

$$p\left(\{v_1 \dots v_{i-1} v_{i+1} \dots v_n\}\right) \rightarrow V_{\setminus i}$$

This $V_{\setminus i}$ is passed through a regression function g to get predicted vector representation v_i' for the predicted block

$$g_\phi\left(v_{\setminus i}\right) \rightarrow v_i'$$

where,
g_ϕ : a neural network with parameters ϕ
ϕ is trainable parameter.

Let q be a function that takes in v_i' and converts this vector back into image space.

$$q_\omega(v_i') \rightarrow I_i'$$

where,

q_ω : a upsampling neural network with parameters ω

is trainable parameter. This generates a block of the predicted Image I. On passing all the block vectors $v_1, v_2, ...v_n$ through d we would get the predicted image I'

Hence when each patch or block v_i tries to predict the original patch x_i Loss in image space can be defined as

$$L_i = \frac{1}{2}\left(I_i' - I_i\right)^2$$

The final objective is defined as

$$J_{\theta,\omega,\phi} = \frac{1}{2}\sum_i \left(q_\omega(g_\phi(p(\{f_\theta(x_1)..f_\theta(x_i-1), f_\theta(x_i+1)..f_\theta(x_n)\}))) - I_i\right)^2$$

2 Region of Interest

- for each block i in x_1,x_2...x_n, r_i is the residuals in vector space, then for higher residuals patch x_i is region of interest.

$$r_i = \frac{1}{2}\left(v_i' - v_i\right)^2$$