SegError Metric Formulae

General Definitions

$$S \coloneqq \text{a partition of the voxels into disjoint subsets } \{S_1, S_2, ..., S_m\}$$
$$T \coloneqq \text{a partition of the voxels into disjoint subsets } \{T_1, T_2, ..., T_l\}$$
$$(T \text{ is the ground truth segment if applicable})$$

$$N \coloneqq \#\text{ voxels in } S = \#\text{ voxels in } T$$

$$s_i \coloneqq |S_i| \qquad t_j \coloneqq |T_j| \qquad c_{ij} \coloneqq |S_i \cap T_j|$$

$$s'_i \coloneqq \frac{s_i}{N} \qquad t'_j \coloneqq \frac{t_j}{N} \qquad p_{ij} \coloneqq \frac{c_{ij}}{N}$$

S and T represent the two segmentations we feed into the error module. Throughout this document, **score** metrics refer to those which increase as the segmentations become more similar, and **error** metrics refer to those which do the opposite.

For quick references to the relevant metric equations, each of these has been numbered.

1 Rand Index-Based Metrics

Let,

$$TP = \#\text{ voxel pairs mapped to the } \textbf{same} \text{ segment within } \textbf{both } S \text{ and } T$$
$$FP = \#\text{ voxel pairs mapped to the } \textbf{same} \text{ segment within } S \text{ but not } T$$
$$FN = \#\text{ voxel pairs mapped to the } \textbf{same} \text{ segment of } T\text{, but not } S$$
$$TN = \#\text{ voxel pairs mapped to the } \textbf{different} \text{ segments of } \textbf{both } S \text{ and } T$$

There are two general ways we can derive Rand Index-Based Metrics for segmentaiton. One involves only counting **distinct** pairs of voxels, and the other also includes counting a voxel by itself.

1.1 Counting Distinct Pairs

Recall the number of distinct pairs within a set of size $n = \binom{n}{2}$. Following this reasoning, we can note that $TP = \sum_{i,j} \binom{c_{ij}}{2}$, $TP + FP = \sum_i \binom{s_i}{2}$, $TP + FN = \sum_j \binom{t_j}{2}$. We can see this for $TP+FN$ by noting that each pair within a segment t_j can either be grouped together by S or not. If they are grouped together, then this pair is a TP pair by the definition above, and if not, then it's a FN pair. Similar reasoning follows for the other two quantities.

Using the forms above, we can form the **Rand Index** (RI) by normalizing the number of true positives and negatives over the total number of pairs.

$$RI = \frac{TP + TN}{\binom{N}{2}} = RS \tag{1}$$

This is returned by the **Rand Score** option (RI or RS) of SegError.

In practice, this value is currently computed in relation to the **Rand Error** (RE), which we can relate by a subtraction from 1.

$$\begin{aligned} RE &= 1 - RS \\ &= \frac{FP + FN}{\binom{N}{2}} \\ &= \frac{(TP + FP) + (TP + FN) - 2TP}{\binom{N}{2}} \end{aligned}$$

$$RE = \frac{\sum_i \binom{s_i}{2} + \sum_j \binom{t_j}{2} - 2\sum_{i,j} \binom{c_{ij}}{2}}{\binom{N}{2}} \tag{2}$$

We commonly decompose different metrics into two versions, one referring to the number of *splits* a segmentation S makes (relative to T), and the other refers to *mergers*.

$$RE_{\text{split}} = \frac{FN}{\binom{N}{2}} = \frac{\sum_j \binom{t_j}{2} - \sum_{i,j} \binom{c_{ij}}{2}}{\binom{N}{2}} \text{ , } RE_{\text{merge}} = \frac{FP}{\binom{N}{2}} = \frac{\sum_i \binom{s_i}{2} - \sum_{i,j} \binom{c_{ij}}{2}}{\binom{N}{2}} \tag{3}$$

1.2 Non-distinct pairs

By now allowing a voxel to form a pair with itself, the number of pairs for a group of size $n = n^2$. Then, $TP = \sum_{i,j} c_{ij}^2$, $TP + FP = \sum_i s_i^2$, $TP + FN = \sum_j t_j^2$, and the respective versions of the equations above take the following form. Note that this takes a condensed alternate form if we instead use the values divided by N (i.e. s_i', t_j', and p_{ij}).

$$sRE = \frac{\sum_i s_i^2 + \sum_j t_j^2 - 2\sum_{i,j} c_{ij}^2}{N^2} = \sum_i s_i'^2 + \sum_j t_j'^2 - \sum_{i,j} p_{i,j}^2$$

$$sRE_{\text{split}} = \sum_j t_j'^2 - \sum_{i,j} p_{ij}^2 \qquad sRE_{\text{merge}} = \sum_i s_i'^2 - \sum_{i,j} p_{ij}^2$$

1.3 Rand F-Score

We can take the version of this metric involving non-distinct pairs, and normalize using either s_i or t_j in order to similarly capture the decomposition between *split* and *merge* scores. This underlies the official error metric for the ISBI2012 competition (`http://brainiac2.mit.edu/isbi_challenge/home`), and is returned by the **Rand F-Score** (RFS) option of SegError, along with its counterpart error (RFE).

$$RFS_{\text{split}} = \frac{\sum_{i,j} p_{ij}^2}{\sum_j t_j'^2} \qquad RFS_{\text{merge}} = \frac{\sum_{i,j} p_{ij}^2}{\sum_i s_i'^2} \tag{4}$$

$$RFS = \frac{\sum_{i,j} p_{ij}^2}{\alpha \sum_i s_i'^2 + (1-\alpha) \sum_j t_j'^2} \tag{5}$$

$$RFE = 1 - RFS \tag{6}$$

1.4 Metric Relations

It is important to note that the components of the Rand F-Score are equivalent to versions of the RE counterparts with different normalization (following the choice of whether or not to count non-distinct pairs). We show the derivation below for the *split* component, where the same results follow for the *merge* component.

$$\begin{aligned}
RE_{\text{split}} &= \frac{FN}{\binom{N}{2}} \\
&= \frac{\sum_j \binom{t_j}{2} - \sum_{i,j} \binom{c_{ij}}{2}}{\binom{N}{2}} \\
&= \frac{\sum_j (t_j^2 - t_j) - \sum_{i,j} (c_{ij}^2 - c_{ij})}{N(N-1)} \\
&= \frac{\sum_j (t_j^2) - n - \sum_{i,j} (c_{ij}^2) + n}{N(N-1)} \\
&= \frac{\sum_j t_j^2 - \sum_{i,j} c_{ij}^2}{N(N-1)} \\
&= \frac{\sum_j t_j^2 - \sum_{i,j} c_{ij}^2}{\sum_j t_j^2} \frac{\sum_j t_j^2}{N(N-1)} \\
&= \frac{\sum_j t_j'^2 - \sum_{i,j} p_{ij}^2}{\sum_j t_j'^2} \frac{\sum_j t_j^2}{N(N-1)} \\
&= (1 - \frac{\sum_{i,j} p_{ij}^2}{\sum_j t_j'^2}) \frac{\sum_j t_j^2}{N(N-1)} \\
&= (1 - RFS_{\text{split}}) \frac{\sum_j t_j^2}{N(N-1)}
\end{aligned}$$

As referred to above, a similar process follows for the *merge* component

$$RE_{\text{merge}} = (1 - RFS_{\text{merge}})\frac{\sum_i s_i^2}{N(N-1)}$$

2 Precision and Recall

We can also explicitly return Precision (PR/sPR) and Recall ($REC/sREC$) as commonly defined, with 's' versions denoting that we allow self loops.

$$PR = \frac{TP}{TP+FP} = \frac{\sum_{i,j}\binom{c_{ij}}{2}}{\sum_i \binom{s_i}{2}} \qquad REC = \frac{TP}{TP+FN} = \frac{\sum_{i,j}\binom{c_{ij}}{2}}{\sum_j \binom{t_j}{2}} \tag{7}$$

$$sPR = \frac{TP}{TP+FP} = \frac{\sum_{i,j} p_{ij}^2}{\sum_i s_i'^2} \qquad sREC = \frac{TP}{TP+FN} = \frac{\sum_{i,j} p_{ij}^2}{\sum_j t_j'^2} \tag{8}$$

3 Information Theory Based Metrics

We also use distance metrics for segmentations based on information theory.

Specifically, given two segmentations S and T, let,

$$H(S) = -\sum_i s_i' \log s_i' \qquad H(T) = -\sum_j t_j' \log t_j'$$

$$H(S|T) = -\sum_{i,j} p_{ij} \log \frac{p_{ij}}{t_j'} \qquad H(T|S) = -\sum_{i,j} p_{ij} \log \frac{p_{ij}}{s_i'}$$

$$I(S,T) = \sum_{i,j} p_{ij} \log \frac{p_{ij}}{s_i' t_j'}$$

$H(S)$ is known as the **Shannon Entropy** for a segmentation S, and $I(S,T)$ to be the **Mutual Information** between S and T. $H(S|T)$ is the **Conditional Entropy** of S given T, which represents the entropy leftover within S after observing a particular value of T.

3.1 Variation of Information

We then calculate the **Variation of Information** (VI or VIE) as

$$VI(S,T) = H(S) + H(T) - 2I(S,T) = H(S|T) + H(T|S)$$

where we use the second form (made explicit below) for our current computation

$$VI = -\sum_{i,j} p_{ij}(\log \frac{p_{ij}}{s_i'} + \log \frac{p_{ij}}{t_j'}) \tag{9}$$

$$VI_{\text{split}} = H(S|T) = -\sum_{i,j} p_{ij} \log \frac{p_{ij}}{t_j'} \qquad VI_{\text{merge}} = H(T|S) = -\sum_{i,j} p_{ij} \log \frac{p_{ij}}{s_i'} \tag{10}$$

This is an unnormalized metric, and so computing the score version (VIS) just returns the inverse.

$$VIS = -1 * VI \tag{11}$$

3.2 Variation F Score

However, we also form a normalized metric for each of the above, and derive a similar **Variation F-score** ($VIFS$) as with the Rand Error. Note that $H(S) = I(S,T) + H(S|T)$

$$VIFS_{\text{split}} = 1 - \frac{H(S|T)}{H(S)} = \frac{I(S,T)}{H(S)} \qquad VIFS_{\text{merge}} = 1 - \frac{H(T|S)}{H(T)} = \frac{I(S,T)}{H(T)} \tag{12}$$

$$VIFS = \frac{I(S,T)}{\alpha H(T) + (1-\alpha)H(S)} \tag{13}$$