Leukocyte Detection

Given a contour

\[ C = (x_1, y_1), \ldots, (x_n, y_n) \]

These contour points are sometimes called 

"snakes".

We want to classify \( C \) as 'leukocyte' or 'non-leukocyte'.

Our measurement or evidence for the contour \( C \) is \( \mathcal{N}_{iCOV} \) computed on \( C \):

\[ X = \frac{m}{\text{ssn}} \text{, when } m = \frac{1}{n} \sum_{i=1}^{n} V_i \text{, } s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (V_i - m)^2 \]

\( e_{iCOV} \)

In Bayes classification we need the posterior

\[ P(C | X) \propto P(X | C) P(C) \text{ likelihood prior} \]

How do we model likelihood & prior?
Modeling the prior

We believe that if a contour $C$ encompasses a leukocyte, if at least $rn$ smaxels of $C$ are on strong edges, where $r$ is a fraction, i.e. $0 < r < 1$.

Q: How do we define "strong" edges?

Q: How do we get $r$?

Let $\Pi_j = \text{Probability that } j \text{ smaxels of a contour } C \text{ are on strong edges.}$

Then the prior becomes

$$\frac{1}{\text{Leukocyte}} \cdot P(C \text{ encompasses a leukocyte})$$

$$= P(C = 'i') = \sum_{j=0}^{n} \Pi_j$$

Using notation,

$$P(C \text{ is not a leukocyte}) = P(C = 'n') = \sum_{j=0}^{r_n-1} \Pi_j$$

Further, we model $\Pi_j$ as binomially distributed, then

$$\Pi_j = \binom{n}{j} p^j (1-p)^{n-j}, \quad p \text{ being a parameter } 0 < p < 1.$$  

So, $P(C = 'i') = \sum_{j=r_n}^{n} \binom{n}{j} p^j (1-p)^{n-j}$

$$P(C = 'n') = \sum_{j=0}^{r_n-1} \binom{n}{j} p^j (1-p)^{n-j}$$

We can easily check these two probabilities sum to 1.

Q: How do we get $p$?
Q: What are strong edges?

Subjective answer: on a few training images, delineate a number of leukocytes by contour software (e.g., spline models). Measure the average image gradient strength along the contour. Consider an edge strong if its gradient magnitude \( \geq \mu \); else this edge is weak.

Q: How do we get \( \nu \)?

On the training images, for the contours delineating leukocytes, compute the average percentage of snarels on strong edges. This could be just one way.

Q: How do we get \( \rho \)?

For a set of training contours. Count the number of snarels sitting on strong edges. Then feed these counts to an ML (maximum likelihood) estimator to learn the parameter \( \rho \) of the binomial distribution.
Modeling Likelihood

\[ P(x|C='n') = \frac{\sum_{j=0}^{n-1} \pi_j P(x|j)}{\sum_{j=0}^{n-1} \pi_j} \]

\[ P(x|C='l') = \frac{\sum_{j=\pi_{n-1}}^{n} \pi_j P(x|j)}{\sum_{j=\pi_{n}}^{n} \pi_j} \]

These are called mixture densities.

\[ P(x|j) = \text{probability density of } x \text{ given that } j \text{ orsels of } C \text{ are on strong edges}. \]

Remember that \[ x = \frac{m}{\sqrt{n}} \]
\[ m = \frac{1}{n} \sum_{i=1}^{n} V_i \]
\[ s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (V_i - m)^2 \]
\[ V_i = \nabla I(x_i, y_i) \cdot \vec{n}(x_i, y_i) \]

Can \( P(x|j) \) be represented in closed form?

Let's see:

\[ V_i \sim N(\mu, \sigma^2) \text{ if } (x_i, y_i) \text{ on a strong edge} \]
\[ V_i \sim N(0, \sigma^2) \text{ if } (x_i, y_i) \text{ on a weak edge} . \]

From training images, we can estimate \( \mu \) and \( \sigma^2 \). These are signal processing techniques.
Next, we can derive

\[
P(x|j=0) = \text{Student-t}(\nu, \nu, \mu, \sigma^2).
\]

\[
P(x|j=n) = \text{non-central t}(\frac{\mu}{\sigma^2}, \nu, \nu, \mu, \sigma^2).
\]

for \(1 \leq j \leq n-1\)

\[
P(x|j) = N(\mu^j, \sigma^j)
\]

\(\sim\) Gaussian.

where \(\mu^j\) and \(\sigma^j\) are functions of \(j, \nu, r, n, \sigma, \mu\).

These derivations are rather painstaking and omitted here. For the details look at [Dawid-Ray-Ackoff] paper.

Once we learn the likelihood and the prior, applying it to a test context is trivial.

\[
C = \begin{cases} \text{'l'} & \text{if } \frac{P(C = \text{'l'}) P(x|C = \text{'l'})}{P(C = \text{'n'}) P(x|C = \text{'n'})} > \frac{\prod_{j=0}^{n-1} P(x|j)}{\prod_{j=1}^{n-1} P(x|j)} \\ \text{'n'} & \text{otherwise} \end{cases}
\]

\[
C = \begin{cases} \text{'l'} & \text{if } \sum_{j=0}^{n} P(x|j) > \sum_{j=1}^{n-1} P(x|j) \\ \text{'n'} & \text{otherwise} \end{cases}
\]