Abstract

Chronic Recurrent Multifocal Osteomyelitis (CRMO) is a rare condition mainly affecting the distal regions of long bones in the body including the femur and tibia [1]. We classify progression of CRMO by considering pairs of MRI images of the knee and long bones of the leg. In this approach, we train multiple classifiers: a logistic classifier with features extracted using a bag of visual words approach [3] for use with pairs of images; we then train SVM with radial basis function and Naive Bayes classifiers on this data.

Models/Methods

We developed two models by extracting different feature sets. Models were trained on a 70/30 training and development split, with K-fold cross validation. A voting ensemble was used to combine the two methods as shown in figure 2.

Transfer Learning

Deep learning is the current state of the art for image classification. Pre-trained models enable quality feature generation where training a custom model is infeasible due to small data.

CNN Feature Extraction

- Pre-processed images to have size 299x299x3.
- Features extracted from final layer of Inception-v3 convolutional neural network [2].
- Low \( \Sigma^2 \) feature reduction.
- Pairs of images represented as the difference between image feature vectors (399 features).

Softmax with Regularization

- L2 regularization to prevent overfitting.
- Cross-entropy cost function:

\[
J(\theta) = \frac{1}{m} \sum_{i=1}^{m} \sum_{k=1}^{n} y_{ik} \log h_\theta(x^{(i)}) + \lambda ||\theta||^2_2
\]

Bag of Visual Words

We modify a bag of visual words approach [3] for use with pairs of images; we then train SVM with radial basis function and Naive Bayes classifiers on this data.

Bag of Visual Words

- Using ORB, SURF, and SIFT, extract features for each image.
- Build visual vocabulary by clustering on the union of features across all images:

\[
V = \{ \mu_1, \ldots, \mu_n \}
\]

where the resulting visual words are obtained via running k-means updates to convergence:

\[
\mu^{(t)} := \arg \min \mu \sum_{i=1}^{m} \frac{1}{||x^{(i)} - \mu||^2} \left| \frac{1}{m} \sum_{i=1}^{m} \left( x^{(i)} = j \right) \right|
\]

- Represent a pair of images as a vector in \( \mathbb{R}^{2n} \) obtained by concatenating the bag of words representation for each image.

Results

We show results for a subset of models which were relatively high-performing in cross-validation. Final test error is shown for all models for completion.

<table>
<thead>
<tr>
<th>Model</th>
<th>Training Error</th>
<th>Dev Error</th>
<th>Test error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-class</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNN Softmax</td>
<td>0.08</td>
<td>0.31</td>
<td>0.64</td>
</tr>
<tr>
<td>NB-BOW, SIFT, (</td>
<td>N</td>
<td>= 500)</td>
<td>0.0</td>
</tr>
<tr>
<td>Ensembled model</td>
<td>0.05</td>
<td>0.17</td>
<td>0.58</td>
</tr>
<tr>
<td>Binary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNN Logistic</td>
<td>0.05</td>
<td>0.12</td>
<td>0.29</td>
</tr>
<tr>
<td>SVM-BOW, ORB, (</td>
<td>N</td>
<td>= 50)</td>
<td>0.0</td>
</tr>
<tr>
<td>Ensembled model</td>
<td>0.11</td>
<td>0.22</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Discussion and Future Work

Our ensemble approach produced models which were more stable and lower-variance than the constituent models in the binary case. In the multi-class case, however, our models failed to generalize past the training data. In particular, models failed to predict the persisted class, which may be due to lack of data for this class. Noisy data and small dataset size made it difficult for models to learn useful information without succumbing to high variance. Visual inspection of incorrectly-classified examples also suggests potential human error in curation. We plan to increase robustness by combining multiple radiologist assessments in the future. Developing heuristics for key point usefulness could decrease variance by minimizing the feature set. We would like to expand on our transfer learning approach by using models trained specifically on WB-MRI data as opposed to general-purpose image classifiers.

References