Building footprints provide useful visual context for users of digital maps. We demonstrate a method for extracting and symbolizing building footprints on the DeepGlobe Challenge data [1], using a convolutional neural network (CNN) with the following characteristics:

- First six (6) layers are the same as VGG-16
- One convolutional layer for detecting building presence
- One convolutional layer for rotated rectangle parameters

In this way, we are able to directly predict rotated rectangles from imagery. This approach works best in suburban areas of North America, such as Las Vegas.

Each best-fitting rotated rectangle is described by 5 parameters: its center \(x, y\), its width and height \(w, h\), and its angle with respect to a horizontal line \(\alpha\). These parameters are relative to a grid cell characterized by a center \(x_g, y_g\) and dimensions \(h_g, w_g\). The angle exists in the range \([-\pi/2, \pi/2]\), and we project it onto the unit circle using its cosine and sine representation. Thus, rotated rectangles are then represented by the following 6 parameters given some grid \(g\):

\[
\begin{align*}
\hat{x} &= x - x_g, \\
\hat{y} &= y - y_g, \\
\hat{h} &= \log h/h_g, \\
\hat{w} &= \log w/w_g, \\
\hat{\alpha} &= \cos 2\alpha, \\
\hat{\alpha}_s &= \sin 2\alpha.
\end{align*}
\]

The proposed approach provides good approximations of small and well-separated buildings which are dominant in U.S. cities. However, when buildings are close to each other, of larger size, or when they have non-rectangular shapes the proposed approach does not allow us to capture them with a reasonable IOU, explaining lower F1 scores in Khartoum and Shanghai.

<table>
<thead>
<tr>
<th>AOI</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Las Vegas</td>
<td>0.733</td>
<td>0.593</td>
<td>0.664</td>
</tr>
<tr>
<td>Paris</td>
<td>0.353</td>
<td>0.258</td>
<td>0.295</td>
</tr>
<tr>
<td>Khartoum</td>
<td>0.243</td>
<td>0.161</td>
<td>0.194</td>
</tr>
<tr>
<td>Shanghai</td>
<td>0.125</td>
<td>0.082</td>
<td>0.099</td>
</tr>
<tr>
<td>Total</td>
<td>0.498</td>
<td>0.365</td>
<td>0.431</td>
</tr>
</tbody>
</table>

The CNN architecture outputs rotated rectangles providing a symbolized approximation for small buildings.

Our CNN architecture performs best on somewhat large, well-separated buildings such as individual houses in a suburban area.

Our method performs best on somewhat large, well-separated buildings which are dominant in U.S. cities. However, when buildings are close to each other, of larger size, or when they have non-rectangular shapes the proposed approach does not allow us to capture them with a reasonable IOU, explaining lower F1 scores in Khartoum and Shanghai.

The proposed approach provides good approximations of small and well-separated buildings which are dominant in U.S. cities. However, when buildings are close to each other, of larger size, or when they have non-rectangular shapes the proposed approach does not allow us to capture them with a reasonable IOU, explaining lower F1 scores in Khartoum and Shanghai.

The CNN architecture outputs rotated rectangles providing a symbolized approximation for small buildings.

Our CNN architecture performs best on somewhat large, well-separated buildings such as individual houses in a suburban area.