Video Coding for Machines: Large-Scale Evaluation of DNNs
Robustness to Compression Artifacts for Semantic Segmentation

Alban Marie\(^1\)  Karol Desnos\(^1\)  Luce Morin\(^1\)  Lu Zhang\(^1\)
\(^1\) Univ Rennes, INSA Rennes, CNRS, IETR - UMR6164; Rennes, France

Example of a Video Coding for Machines Pipeline

*Low computation on camera side.*
*Low data transmission to the vision task side.*
*Degradation introduced by lossy compression.*

How resilient a semantic segmentation algorithm is to various compression artifacts?

Considered Coding Configurations

Which encoding strategy should be followed to achieve optimal bitrate accuracy trade-off?

A total of 1486 coding configurations are considered, including:

- ✔ Images with or without colors
- ✔ Wide range of image resolutions
- ✔ 5 codecs from JPEG to VVenC
- ✔ Wide range of bitrates

Progressive Training Procedure

- ✗ A vision task trained on pristine images performs poorly on distorted content.
- ✔ Mitigated by re-training with distorted images using the proposed progressive training:
  \[ f(e) = p_\infty + \Delta p \left[ \frac{1}{\Delta p} (p_0 - p_\infty) \exp(-se) \right] \]
- ✔ Progressive training allows to re-train one DNN on a large amount of distortion at once, ranging from undistorted to highly distorted.

Experimental Results

- ✔ 58.3%, 49.8%, 33.5% and 24.3% bitrate savings with optimal image resolution for JPG, JM, x265 and VVenC, respectively.
- ✔ BDR increase when removing chrominance channels, except at very low rates for JPG/J2K
- ✔ JPG can significantly outperform VTM with re-training and optimal image resolution

Comparison of the proposed progressive training with other training strategies.

<table>
<thead>
<tr>
<th>Complexity↓</th>
<th>BDR↓</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>—</td>
</tr>
<tr>
<td>Separate train</td>
<td>100.00%</td>
</tr>
<tr>
<td>Data augm. [1]</td>
<td>46.99%</td>
</tr>
<tr>
<td>ours, s = 0.050</td>
<td>26.51%</td>
</tr>
<tr>
<td>ours, s = 0.045</td>
<td>48.19%</td>
</tr>
<tr>
<td>ours, s = 0.025</td>
<td>80.75%</td>
</tr>
</tbody>
</table>