Combining GANs and AutoEncoders for Efficient Anomaly Detection

Fabio Carrara\textsuperscript{1}, Giuseppe Amato\textsuperscript{1}, Luca Brombin\textsuperscript{2}, Fabrizio Falchi\textsuperscript{1}, Claudio Gennaro\textsuperscript{1}

\textsuperscript{1}ISTI CNR, Pisa, Italy — \textsuperscript{2}University of Pisa

fabio.carrara@isti.cnr.it

https://github.com/fabiocarrara/cbigan-adv

ICPR 2021 - January, 14th
Milan, Italy (Virtual)
The Problem: Visual Anomaly Detection

**GOAL** — detect anomalous images (assign either ✓ or ✘ to input image)

**HOW** — Fit *generative models* to reconstruct non-anomalous inputs (one-class training, only ✓ are commonly available in the training phase)

- Small difference w.r.t input ⇒ ✓
- Larger difference w.r.t input (defects are missing) ⇒ ✘
Generative models in reconstruction-based AD often belong to **GAN** or **AE** families.

Two main classes of methodology:

- **Iterative methods** — reconstructions are optimized via multiple iterations for each input
  - better reconstructions, but very expensive
  - Es: AnoGAN, VAE-grad

- **Single-pass methods** — reconstruction obtained in a single forward pass (often an Encoder-Decoder)
  - more efficient, less precise reconstructions
  - Es: EGBAD (BiGAN), AE\_L2/SSIM, AVID, LSA, etc.

Among single-pass methods, the two commonly approaches adopted are:

- **GANs**
  - + realistic outputs
  - - (s)low control on reconstruction

- **AutoEncoders**
  - + preserve alignment w.r.t input
  - - blurry reconstructions
CBiGAN: Consistency BiGAN Model

We propose CBiGAN — a combined model.

+ **generalizes** both Bidirectional GANs ($\alpha = 0$) and AutoEncoders ($\alpha = 1$)

+ $\alpha$ can be tuned to balance the regularization

+ produces **realistic** outputs **fast** that are **consistent** with inputs

\[
\mathcal{L}_{\text{total}} = (1 - \alpha) \cdot \mathcal{L}_{\text{adversarial}} + \alpha \cdot \mathcal{L}_{\text{consistency}}
\]

pushes $\text{Gen}$ and $\text{Enc}$ to make realistic results, $\text{Discr}$ to spot fakes

pushes $\text{Gen}$ and $\text{Enc}$ to produce outputs consistent with inputs
**Evaluation: MVTec-AD Dataset**

**MVTec-AD — benchmark of visual AD for industrial applications**

<table>
<thead>
<tr>
<th>Methods</th>
<th>Textures</th>
<th>Objects</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Iterative Methods</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AnoGAN</td>
<td>0.54</td>
<td>0.56</td>
<td>0.55</td>
</tr>
<tr>
<td>VAE-grad</td>
<td>0.78</td>
<td><strong>0.76</strong></td>
<td><strong>0.77</strong></td>
</tr>
<tr>
<td><strong>Single-pass Methods</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AE_L2</td>
<td>0.65</td>
<td>0.74</td>
<td>0.71</td>
</tr>
<tr>
<td>AVID</td>
<td>0.67</td>
<td><strong>0.75</strong></td>
<td>0.73</td>
</tr>
<tr>
<td>LSA</td>
<td>0.69</td>
<td>0.75</td>
<td>0.73</td>
</tr>
<tr>
<td>EGBAD (BiGAN)</td>
<td>0.66</td>
<td>0.58</td>
<td>0.61</td>
</tr>
<tr>
<td><strong>CBiGAN (ours)</strong></td>
<td><strong>0.84</strong></td>
<td>0.73</td>
<td>0.76</td>
</tr>
</tbody>
</table>

↑ Balanced Accuracy = (TPR + FPR) / 2 when using the Youden threshold. Best among single-pass methods.

↓ Area Under the ROC Curve (AuROC)

<table>
<thead>
<tr>
<th>Methods</th>
<th>Textures</th>
<th>Objects</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE_L2</td>
<td>0.80</td>
<td>0.74</td>
<td>0.75</td>
</tr>
<tr>
<td>GeoTrans</td>
<td>0.59</td>
<td>0.71</td>
<td>0.67</td>
</tr>
<tr>
<td>GANomaly</td>
<td>0.77</td>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>EGBAD</td>
<td>0.66</td>
<td>0.57</td>
<td>0.60</td>
</tr>
<tr>
<td><strong>CBiGAN (ours)</strong></td>
<td><strong>0.85</strong></td>
<td>0.73</td>
<td>0.77</td>
</tr>
</tbody>
</table>
Conclusions and Future Work

- **GOAL** — detect anomalous images: assign either ✓ or ✘ to input image
- **CONSTRAINTS** — one-class AD: only good images available
- **METHOD** — reconstruction-based AD: difference between input and reconstruction
  - Iterative vs Single-Pass methods: effectiveness vs efficiency
- **CONTRIBUTION** — we propose Consistency BiGAN for one-class single-pass AD
  - combines and generalizes Bidirectional GANs and AutoEncoders
  - provides reconstruction ability of the former and consistency of the latter
  - efficient (single-pass) with results comparable to iterative methods
- **FUTURE WORK** — address Anomaly Localization (assign either ✓ or ✘ to each pixel)