Exploring the Current State of the Art in Deep Multi-View-Clustering
Michael Oster & Alexander Sturm supervised by Lukas Miklautz & Sebastian Tschiatschek

Introduction and contributions

- Multi-View-Clustering (MVC) generalizes the clustering setting to multi-view data
- E.g. images of same object taken from different angles
- Recently Trosten et al. invented a new approach to MVC called AECoDDC, which uses:
  - reconstruction loss as single-view self-supervised-learning objective (SV-SSL)
  - contrastive loss as the multi-view self-supervised-learning objective (MV-SSL)
  - Deep Divergence-based (DDC) clustering loss for the Clustering module

Contributions of our project:
- developed new architectures building upon AECoDDC
- investigated how methods using contrastive alignment suffer from noise
- looked at the benefits of splitting the latent space into cluster-relevant and -irrelevant information
- conducted extensive latent space exploration

Architectures

- Introduce ACe/DeC module after training to split into cluster-relevant and -irrelevant
- Same behaviour as AECoDDC
- Include ACe/DeC module within original architecture
- New Compression loss term while training leads to different behaviour

Results for “Noisy Fashion” MVC dataset

PCA plot of latent space

<table>
<thead>
<tr>
<th>Architecture</th>
<th>ACC</th>
<th>NMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AECoDDC (prior best)</td>
<td>0.8 (0.02)</td>
<td>0.77 (0.01)</td>
</tr>
<tr>
<td>AECoCoDDC</td>
<td><strong>0.86 (0.02)</strong></td>
<td><strong>0.81 (0.01)</strong></td>
</tr>
</tbody>
</table>

Conclusion:
- Contrastive learning benefits from ground truth label information embedded in the views
- AECoDDC+ACe/DeC is able to split into cluster-relevant and -irrelevant
- The model finds human interpretable centroids
- Performance improvement on “Noisy Fashion” does not generalize to more complex datasets

References:
- Lukas Miklautz et al.: “Details (Don’t) Matter: Isolating Cluster Information in Deep Embedded Spaces”