Introduction to machine learning and deep learning

Convolutional Neural Networks

Presented by Benoit Liquet
University of Pau et des Pays de l'Adour

1. Introduction to Machine Learning and Deep Learning
   1.1 Introduction
   1.2 Supervised Machine Learning: Overview
   1.3 Two Fundamental Operations
   1.4 Training a Neural Network
   1.5 Building a Convolutional Layer
      - Can we see what convolutional networks do "under the hood"?
      - Convolutional Neural Networks: Overview
      - Convolutional Neural Networks: Details
      - Receptive fields, Convolution, Padding, Stride, and Bias
      - Issues with Multiple Channels
      - Pooling and Convolutional Layers
   1.6 Building a Convolutional Neural Network
      - Convolutional Neural Networks: Overview
      - Pooling Layers
      - Flattening Layers
      - Fully Connected Layers
      - Putting it all together
      - Backpropagation
      - Learnable Hyperparameters
      - Learnable Weight Layers and Channel Features
      - Learnable Bias Layers
      - Activation Layers
   1.7 Regularization
      - Dropout Layers
      - Batch Normalization Layers and Other Regularization Layers
      - Data Augmentation
   1.8 Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.9 Modern Practical Advances
   1.10 Conclusion: Beyond Convolutional Neural Networks
   1.11 Appendix: Mathematics Behind Convolutional Neural Networks
   1.12 Appendix: Proofs and Derivations
   1.13 Appendix: Practical Advanced Techniques
   1.14 Appendix: Pooling Layers Backpropagation
   1.15 Appendix: Batch Normalization Backpropagation
   1.16 Appendix: Dropout Backpropagation
   1.17 Appendix: Convolutional Layers and Channel Features
   1.18 Appendix: Bias Backpropagation
   1.19 Appendix: Activation Backpropagation
   1.20 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.21 Appendix: Modern Practical Advances
   1.22 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.23 Appendix: Convolutional Layers and Channel Features
   1.24 Appendix: Bias Backpropagation
   1.25 Appendix: Activation Backpropagation
   1.26 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.27 Appendix: Modern Practical Advances
   1.28 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.29 Appendix: Convolutional Layers and Channel Features
   1.30 Appendix: Bias Backpropagation
   1.31 Appendix: Activation Backpropagation
   1.32 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.33 Appendix: Modern Practical Advances
   1.34 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.35 Appendix: Convolutional Layers and Channel Features
   1.36 Appendix: Bias Backpropagation
   1.37 Appendix: Activation Backpropagation
   1.38 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.39 Appendix: Modern Practical Advances
   1.40 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.41 Appendix: Convolutional Layers and Channel Features
   1.42 Appendix: Bias Backpropagation
   1.43 Appendix: Activation Backpropagation
   1.44 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.45 Appendix: Modern Practical Advances
   1.46 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.47 Appendix: Convolutional Layers and Channel Features
   1.48 Appendix: Bias Backpropagation
   1.49 Appendix: Activation Backpropagation
   1.50 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.51 Appendix: Modern Practical Advances
   1.52 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.53 Appendix: Convolutional Layers and Channel Features
   1.54 Appendix: Bias Backpropagation
   1.55 Appendix: Activation Backpropagation
   1.56 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.57 Appendix: Modern Practical Advances
   1.58 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.59 Appendix: Convolutional Layers and Channel Features
   1.60 Appendix: Bias Backpropagation
   1.61 Appendix: Activation Backpropagation
   1.62 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.63 Appendix: Modern Practical Advances
   1.64 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.65 Appendix: Convolutional Layers and Channel Features
   1.66 Appendix: Bias Backpropagation
   1.67 Appendix: Activation Backpropagation
   1.68 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.69 Appendix: Modern Practical Advances
   1.70 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.71 Appendix: Convolutional Layers and Channel Features
   1.72 Appendix: Bias Backpropagation
   1.73 Appendix: Activation Backpropagation
   1.74 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.75 Appendix: Modern Practical Advances
   1.76 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.77 Appendix: Convolutional Layers and Channel Features
   1.78 Appendix: Bias Backpropagation
   1.79 Appendix: Activation Backpropagation
   1.80 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.81 Appendix: Modern Practical Advances
   1.82 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.83 Appendix: Convolutional Layers and Channel Features
   1.84 Appendix: Bias Backpropagation
   1.85 Appendix: Activation Backpropagation
   1.86 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.87 Appendix: Modern Practical Advances
   1.88 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.89 Appendix: Convolutional Layers and Channel Features
   1.90 Appendix: Bias Backpropagation
   1.91 Appendix: Activation Backpropagation
   1.92 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.93 Appendix: Modern Practical Advances
   1.94 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.95 Appendix: Convolutional Layers and Channel Features
   1.96 Appendix: Bias Backpropagation
   1.97 Appendix: Activation Backpropagation
   1.98 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.99 Appendix: Modern Practical Advances
   1.100 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.101 Appendix: Convolutional Layers and Channel Features
   1.102 Appendix: Bias Backpropagation
   1.103 Appendix: Activation Backpropagation
   1.104 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.105 Appendix: Modern Practical Advances
   1.106 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.107 Appendix: Convolutional Layers and Channel Features
   1.108 Appendix: Bias Backpropagation
   1.109 Appendix: Activation Backpropagation
   1.110 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.111 Appendix: Modern Practical Advances
   1.112 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.113 Appendix: Convolutional Layers and Channel Features
   1.114 Appendix: Bias Backpropagation
   1.115 Appendix: Activation Backpropagation
   1.116 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.117 Appendix: Modern Practical Advances
   1.118 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.119 Appendix: Convolutional Layers and Channel Features
   1.120 Appendix: Bias Backpropagation
   1.121 Appendix: Activation Backpropagation
   1.122 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.123 Appendix: Modern Practical Advances
   1.124 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.125 Appendix: Convolutional Layers and Channel Features
   1.126 Appendix: Bias Backpropagation
   1.127 Appendix: Activation Backpropagation
   1.128 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.129 Appendix: Modern Practical Advances
   1.130 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.131 Appendix: Convolutional Layers and Channel Features
   1.132 Appendix: Bias Backpropagation
   1.133 Appendix: Activation Backpropagation
   1.134 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.135 Appendix: Modern Practical Advances
   1.136 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.137 Appendix: Convolutional Layers and Channel Features
   1.138 Appendix: Bias Backpropagation
   1.139 Appendix: Activation Backpropagation
   1.140 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.141 Appendix: Modern Practical Advances
   1.142 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.143 Appendix: Convolutional Layers and Channel Features
   1.144 Appendix: Bias Backpropagation
   1.145 Appendix: Activation Backpropagation
   1.146 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.147 Appendix: Modern Practical Advances
   1.148 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.149 Appendix: Convolutional Layers and Channel Features
   1.150 Appendix: Bias Backpropagation
   1.151 Appendix: Activation Backpropagation
   1.152 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.153 Appendix: Modern Practical Advances
   1.154 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.155 Appendix: Convolutional Layers and Channel Features
   1.156 Appendix: Bias Backpropagation
   1.157 Appendix: Activation Backpropagation
   1.158 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.159 Appendix: Modern Practical Advances
   1.160 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.161 Appendix: Convolutional Layers and Channel Features
   1.162 Appendix: Bias Backpropagation
   1.163 Appendix: Activation Backpropagation
   1.164 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.165 Appendix: Modern Practical Advances
   1.166 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.167 Appendix: Convolutional Layers and Channel Features
   1.168 Appendix: Bias Backpropagation
   1.169 Appendix: Activation Backpropagation
   1.170 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.171 Appendix: Modern Practical Advances
   1.172 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.173 Appendix: Convolutional Layers and Channel Features
   1.174 Appendix: Bias Backpropagation
   1.175 Appendix: Activation Backpropagation
   1.176 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.177 Appendix: Modern Practical Advances
   1.178 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.179 Appendix: Convolutional Layers and Channel Features
   1.180 Appendix: Bias Backpropagation
   1.181 Appendix: Activation Backpropagation
   1.182 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.183 Appendix: Modern Practical Advances
   1.184 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.185 Appendix: Convolutional Layers and Channel Features
   1.186 Appendix: Bias Backpropagation
   1.187 Appendix: Activation Backpropagation
   1.188 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.189 Appendix: Modern Practical Advances
   1.190 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.191 Appendix: Convolutional Layers and Channel Features
   1.192 Appendix: Bias Backpropagation
   1.193 Appendix: Activation Backpropagation
   1.194 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.195 Appendix: Modern Practical Advances
   1.196 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.197 Appendix: Convolutional Layers and Channel Features
   1.198 Appendix: Bias Backpropagation
   1.199 Appendix: Activation Backpropagation
   1.200 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.201 Appendix: Modern Practical Advances
   1.202 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.203 Appendix: Convolutional Layers and Channel Features
   1.204 Appendix: Bias Backpropagation
   1.205 Appendix: Activation Backpropagation
   1.206 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.207 Appendix: Modern Practical Advances
   1.208 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.209 Appendix: Convolutional Layers and Channel Features
   1.210 Appendix: Bias Backpropagation
   1.211 Appendix: Activation Backpropagation
   1.212 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.213 Appendix: Modern Practical Advances
   1.214 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.215 Appendix: Convolutional Layers and Channel Features
   1.216 Appendix: Bias Backpropagation
   1.217 Appendix: Activation Backpropagation
   1.218 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.219 Appendix: Modern Practical Advances
   1.220 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.221 Appendix: Convolutional Layers and Channel Features
   1.222 Appendix: Bias Backpropagation
   1.223 Appendix: Activation Backpropagation
   1.224 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.225 Appendix: Modern Practical Advances
   1.226 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.227 Appendix: Convolutional Layers and Channel Features
   1.228 Appendix: Bias Backpropagation
   1.229 Appendix: Activation Backpropagation
   1.230 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.231 Appendix: Modern Practical Advances
   1.232 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.233 Appendix: Convolutional Layers and Channel Features
   1.234 Appendix: Bias Backpropagation
   1.235 Appendix: Activation Backpropagation
   1.236 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.237 Appendix: Modern Practical Advances
   1.238 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.239 Appendix: Convolutional Layers and Channel Features
   1.240 Appendix: Bias Backpropagation
   1.241 Appendix: Activation Backpropagation
   1.242 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.243 Appendix: Modern Practical Advances
   1.244 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.245 Appendix: Convolutional Layers and Channel Features
   1.246 Appendix: Bias Backpropagation
   1.247 Appendix: Activation Backpropagation
   1.248 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.249 Appendix: Modern Practical Advances
   1.250 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.251 Appendix: Convolutional Layers and Channel Features
   1.252 Appendix: Bias Backpropagation
   1.253 Appendix: Activation Backpropagation
   1.254 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.255 Appendix: Modern Practical Advances
   1.256 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.257 Appendix: Convolutional Layers and Channel Features
   1.258 Appendix: Bias Backpropagation
   1.259 Appendix: Activation Backpropagation
   1.260 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.261 Appendix: Modern Practical Advances
   1.262 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.263 Appendix: Convolutional Layers and Channel Features
   1.264 Appendix: Bias Backpropagation
   1.265 Appendix: Activation Backpropagation
   1.266 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.267 Appendix: Modern Practical Advances
   1.268 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.269 Appendix: Convolutional Layers and Channel Features
   1.270 Appendix: Bias Backpropagation
   1.271 Appendix: Activation Backpropagation
   1.272 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.273 Appendix: Modern Practical Advances
   1.274 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.275 Appendix: Convolutional Layers and Channel Features
   1.276 Appendix: Bias Backpropagation
   1.277 Appendix: Activation Backpropagation
   1.278 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.279 Appendix: Modern Practical Advances
   1.280 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.281 Appendix: Convolutional Layers and Channel Features
   1.282 Appendix: Bias Backpropagation
   1.283 Appendix: Activation Backpropagation
   1.284 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.285 Appendix: Modern Practical Advances
   1.286 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.287 Appendix: Convolutional Layers and Channel Features
   1.288 Appendix: Bias Backpropagation
   1.289 Appendix: Activation Backpropagation
   1.290 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.291 Appendix: Modern Practical Advances
   1.292 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.293 Appendix: Convolutional Layers and Channel Features
   1.294 Appendix: Bias Backpropagation
   1.295 Appendix: Activation Backpropagation
   1.296 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.297 Appendix: Modern Practical Advances
   1.298 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.299 Appendix: Convolutional Layers and Channel Features
   1.300 Appendix: Bias Backpropagation
   1.301 Appendix: Activation Backpropagation
   1.302 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.303 Appendix: Modern Practical Advances
   1.304 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.305 Appendix: Convolutional Layers and Channel Features
   1.306 Appendix: Bias Backpropagation
   1.307 Appendix: Activation Backpropagation
   1.308 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.309 Appendix: Modern Practical Advances
   1.310 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.311 Appendix: Convolutional Layers and Channel Features
   1.312 Appendix: Bias Backpropagation
   1.313 Appendix: Activation Backpropagation
   1.314 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.315 Appendix: Modern Practical Advances
   1.316 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.317 Appendix: Convolutional Layers and Channel Features
   1.318 Appendix: Bias Backpropagation
   1.319 Appendix: Activation Backpropagation
   1.320 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.321 Appendix: Modern Practical Advances
   1.322 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.323 Appendix: Convolutional Layers and Channel Features
   1.324 Appendix: Bias Backpropagation
   1.325 Appendix: Activation Backpropagation
   1.326 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.327 Appendix: Modern Practical Advances
   1.328 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.329 Appendix: Convolutional Layers and Channel Features
   1.330 Appendix: Bias Backpropagation
   1.331 Appendix: Activation Backpropagation
   1.332 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.333 Appendix: Modern Practical Advances
   1.334 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.335 Appendix: Convolutional Layers and Channel Features
   1.336 Appendix: Bias Backpropagation
   1.337 Appendix: Activation Backpropagation
   1.338 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.339 Appendix: Modern Practical Advances
   1.340 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.341 Appendix: Convolutional Layers and Channel Features
   1.342 Appendix: Bias Backpropagation
   1.343 Appendix: Activation Backpropagation
   1.344 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.345 Appendix: Modern Practical Advances
   1.346 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.347 Appendix: Convolutional Layers and Channel Features
   1.348 Appendix: Bias Backpropagation
   1.349 Appendix: Activation Backpropagation
   1.350 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.351 Appendix: Modern Practical Advances
   1.352 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.353 Appendix: Convolutional Layers and Channel Features
   1.354 Appendix: Bias Backpropagation
   1.355 Appendix: Activation Backpropagation
   1.356 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.357 Appendix: Modern Practical Advances
   1.358 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.359 Appendix: Convolutional Layers and Channel Features
   1.360 Appendix: Bias Backpropagation
   1.361 Appendix: Activation Backpropagation
   1.362 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.363 Appendix: Modern Practical Advances
   1.364 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.365 Appendix: Convolutional Layers and Channel Features
   1.366 Appendix: Bias Backpropagation
   1.367 Appendix: Activation Backpropagation
   1.368 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.369 Appendix: Modern Practical Advances
   1.370 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.371 Appendix: Convolutional Layers and Channel Features
   1.372 Appendix: Bias Backpropagation
   1.373 Appendix: Activation Backpropagation
   1.374 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.375 Appendix: Modern Practical Advances
   1.376 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.377 Appendix: Convolutional Layers and Channel Features
   1.378 Appendix: Bias Backpropagation
   1.379 Appendix: Activation Backpropagation
   1.380 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.381 Appendix: Modern Practical Advances
   1.382 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.383 Appendix: Convolutional Layers and Channel Features
   1.384 Appendix: Bias Backpropagation
   1.385 Appendix: Activation Backpropagation
   1.386 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.387 Appendix: Modern Practical Advances
   1.388 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.389 Appendix: Convolutional Layers and Channel Features
   1.390 Appendix: Bias Backpropagation
   1.391 Appendix: Activation Backpropagation
   1.392 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.393 Appendix: Modern Practical Advances
   1.394 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.395 Appendix: Convolutional Layers and Channel Features
   1.396 Appendix: Bias Backpropagation
   1.397 Appendix: Activation Backpropagation
   1.398 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.399 Appendix: Modern Practical Advances
   1.400 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.401 Appendix: Convolutional Layers and Channel Features
   1.402 Appendix: Bias Backpropagation
   1.403 Appendix: Activation Backpropagation
   1.404 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.405 Appendix: Modern Practical Advances
   1.406 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.407 Appendix: Convolutional Layers and Channel Features
   1.408 Appendix: Bias Backpropagation
   1.409 Appendix: Activation Backpropagation
   1.410 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.411 Appendix: Modern Practical Advances
   1.412 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.413 Appendix: Convolutional Layers and Channel Features
   1.414 Appendix: Bias Backpropagation
   1.415 Appendix: Activation Backpropagation
   1.416 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.417 Appendix: Modern Practical Advances
   1.418 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.419 Appendix: Convolutional Layers and Channel Features
   1.420 Appendix: Bias Backpropagation
   1.421 Appendix: Activation Backpropagation
   1.422 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.423 Appendix: Modern Practical Advances
   1.424 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.425 Appendix: Convolutional Layers and Channel Features
   1.426 Appendix: Bias Backpropagation
   1.427 Appendix: Activation Backpropagation
   1.428 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.429 Appendix: Modern Practical Advances
   1.430 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.431 Appendix: Convolutional Layers and Channel Features
   1.432 Appendix: Bias Backpropagation
   1.433 Appendix: Activation Backpropagation
   1.434 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.435 Appendix: Modern Practical Advances
   1.436 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.437 Appendix: Convolutional Layers and Channel Features
   1.438 Appendix: Bias Backpropagation
   1.439 Appendix: Activation Backpropagation
   1.440 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.441 Appendix: Modern Practical Advances
   1.442 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.443 Appendix: Convolutional Layers and Channel Features
   1.444 Appendix: Bias Backpropagation
   1.445 Appendix: Activation Backpropagation
   1.446 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.447 Appendix: Modern Practical Advances
   1.448 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.449 Appendix: Convolutional Layers and Channel Features
   1.450 Appendix: Bias Backpropagation
   1.451 Appendix: Activation Backpropagation
   1.452 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.453 Appendix: Modern Practical Advances
   1.454 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.455 Appendix: Convolutional Layers and Channel Features
   1.456 Appendix: Bias Backpropagation
   1.457 Appendix: Activation Backpropagation
   1.458 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.459 Appendix: Modern Practical Advances
   1.460 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.461 Appendix: Convolutional Layers and Channel Features
   1.462 Appendix: Bias Backpropagation
   1.463 Appendix: Activation Backpropagation
   1.464 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.465 Appendix: Modern Practical Advances
   1.466 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.467 Appendix: Convolutional Layers and Channel Features
   1.468 Appendix: Bias Backpropagation
   1.469 Appendix: Activation Backpropagation
   1.470 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.471 Appendix: Modern Practical Advances
   1.472 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.473 Appendix: Convolutional Layers and Channel Features
   1.474 Appendix: Bias Backpropagation
   1.475 Appendix: Activation Backpropagation
   1.476 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.477 Appendix: Modern Practical Advances
   1.478 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.479 Appendix: Convolutional Layers and Channel Features
   1.480 Appendix: Bias Backpropagation
   1.481 Appendix: Activation Backpropagation
   1.482 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.483 Appendix: Modern Practical Advances
   1.484 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.485 Appendix: Convolutional Layers and Channel Features
   1.486 Appendix: Bias Backpropagation
   1.487 Appendix: Activation Backpropagation
   1.488 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.489 Appendix: Modern Practical Advances
   1.490 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.491 Appendix: Convolutional Layers and Channel Features
   1.492 Appendix: Bias Backpropagation
   1.493 Appendix: Activation Backpropagation
   1.494 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.495 Appendix: Modern Practical Advances
   1.496 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.497 Appendix: Convolutional Layers and Channel Features
   1.498 Appendix: Bias Backpropagation
   1.499 Appendix: Activation Backpropagation
   1.500 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.501 Appendix: Modern Practical Advances
   1.502 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.503 Appendix: Convolutional Layers and Channel Features
   1.504 Appendix: Bias Backpropagation
   1.505 Appendix: Activation Backpropagation
   1.506 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.507 Appendix: Modern Practical Advances
   1.508 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.509 Appendix: Convolutional Layers and Channel Features
   1.510 Appendix: Bias Backpropagation
   1.511 Appendix: Activation Backpropagation
   1.512 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.513 Appendix: Modern Practical Advances
   1.514 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.515 Appendix: Convolutional Layers and Channel Features
   1.516 Appendix: Bias Backpropagation
   1.517 Appendix: Activation Backpropagation
   1.518 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.519 Appendix: Modern Practical Advances
   1.520 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.521 Appendix: Convolutional Layers and Channel Features
   1.522 Appendix: Bias Backpropagation
   1.523 Appendix: Activation Backpropagation
   1.524 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.525 Appendix: Modern Practical Advances
   1.526 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.527 Appendix: Convolutional Layers and Channel Features
   1.528 Appendix: Bias Backpropagation
   1.529 Appendix: Activation Backpropagation
   1.530 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.531 Appendix: Modern Practical Advances
   1.532 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.533 Appendix: Convolutional Layers and Channel Features
   1.534 Appendix: Bias Backpropagation
   1.535 Appendix: Activation Backpropagation
   1.536 Appendix: Regularization Techniques and Any Computer Vision Tasks
      - Classification
      - Object Localization
      - Semantic Segmentation
   1.537 Appendix: Modern Practical Advances
   1.538 Appendix: Conclusion: Beyond Convolutional Neural Networks
   1.539 Appendix: Convolutional Layers and Channel

$$L = L_{cls} + L_{loc}$$

- Here L_{cls} counts the number of mistake labels, whereas L_{loc} counts the number of bypass that do not have traceable parameters. Further, the forward loss function consists of two terms: L_{cls} measures the classification loss, and L_{loc} measures the localization loss.
- For example, in the VGG16 network:

$$L = 2k, \quad L_{cls} = 1k, \quad L_{loc} = 1k, \quad L_{cls} + L_{loc} = 2k$$

- See the following figure to verify these numbers:

Figure 6.21: The VGG16 Network Architecture.

- Similar to a feed forward network, the goal of a convolutional neural network is to approximate some unknown function $f^*(.)$. For convolutional neural network, the unknown function $f^*(.)$ is a function that maps an image I to a vector y that gives a probability vector with the highest weight on the class associated with the label of the image I.
  - A convolutional neural network defines a mapping $f(.)$ and learns the values of the unknown parameters θ that ideally result in $f^*(x) \approx f_\theta(x)$ for every input image x in possible.
- In general, similar to feed forward networks, the approximating function $f_\theta(.)$ is recursively composed as:

$$f_\theta(x) = f_{\theta_n}(f_{\theta_{n-1}}(\dots f_{\theta_1}(x)))$$

where for each i, the function $f_{\theta_i}(.)$ is associated with the ith layer which depends on the layer's parameters $\theta^{(i)} \in \theta^{(i)}$.
  - Note that for layers that do not manipulate (or control) on $f_{\theta_i(x)}$ the parameter space $\theta^{(i)}$ is empty.
- Similarly to feed forward networks, it is useful to denote the neuron activations of the network as $a^{(i)}, a^{(i)}, \dots, a^{(n)}$ where $a^{(i)}$ is the output, and for $i = 1, \dots, L - 1$,

$$a^{(i)} = f^{(i)}(a^{(i-1)})$$

with $a^{(0)} = x$.

Convolutional Layers

- When the ith layer of the network is a convolution layer, then $f_{\theta_i}^{(i)}(.)$ is type the output $a^{(i-1)}$ of the previous layer as the input and convolute convolution.
  - In this case the input and output are generally 3-dimensional tensors, as we have seen in the previously.

Pooling Layers

- As mentioned above, there are also non-trainable layers (or control layers) and these are typically called pooling layers. The main idea of a pooling layer is to reduce the height and width of the input tensor $I^{(i-1)}$ to achieve a lower dimensional output $a^{(i)}$ without changing the depth.
- Generally for some fixed channel j and pixel coordinates of the output (i, k), a pooling operation operates on pixels from a window in the input denoted as I_{j,i_0,k_0}. Pooling uses I_{j,i_0,k_0} is a set of pixel coordinates in the input that depends on the output spatial pixel (i, k). There are two common types of pooling layers: max pooling and average pooling.
- For each channel j, the pooling operation can be summarized as:

$$\text{Max}_{j,i,k} = \max_{(i_0,k_0) \in I_{j,i_0,k_0}} I_{j,i_0,k_0} \quad \text{(max pooling)}$$

$$\text{Avg}_{j,i,k} = \frac{1}{|I_{j,i_0,k_0}|} \sum_{(i_0,k_0) \in I_{j,i_0,k_0}} I_{j,i_0,k_0} \quad \text{(average pooling)}$$

- As a resident, max pooling takes the maximal pixel value within the window as the output, while average pooling averages pixel values within the window for the output. See the following figure for an illustration.

Figure 6.11: An example of pooling with a 2×2 window. (a) Max-pooling. (b) Average-pooling.

Note: The idea of pooling integrates with the notion that the **initial layers** of a convolutional network focus on **pixel level features** similar to edge detection, and as we progress towards the final layers of the network, the information is aggregated to address **general questions** about the whole image.

- Thus deeper layers are less sensitive to translation changes on the input image compared to the initial layers.
- For instance, the answer to a question "**Is there a bird in the photo?**" is the same regardless of whether a bird appears in the image, even though the corresponding pixels from the initial layers differ.
- Pooling layers are applied after convolutional layers to support this aggregation by **progressively reducing the spatial dimensions**, allowing the network to recognize larger patterns while focusing less on specific pixel locations.

Fully Connected Layers

- Some layers of a CNN can be fully connected. Such layers are typically deployed at the end of the network. This is because the typical task of a CNN is to predict a probabilistic overall network to answer questions, such as classification of the objects in the image.
- Note that since fully connected layers operate on vectors as the input, in cases the previous layer has a tensor as output, it is flattened to a vector.

Practice Exercise

Practical 1:

- Open Tutorial 2 of CNN available on the workshop GitHub page. Alternatively, click here.
- Save a copy of this in your Google Colab.
- In this exercise, we build a simple CNN for classification using CIFAR10 dataset.

VGG19 Revisited

- We now take a closer look at the architecture of our running example network, VGG19.
- While this is not the most modern convolutional architecture, it is instructive to consider it here since it lists directly within the paradigms we have seen.
- Table below provides complete details.

Layers	Type of Layer	Output Dimension	No. of Neurons	No. of Learned Parameters
1	Input	$3 \times 224 \times 224$	-	-
2	Convolution	$64 \times 224 \times 224$	$3.211.264$	1.792
3	Convolution	$64 \times 224 \times 224$	$3.211.264$	36.928
4	Max-pooling	$64 \times 112 \times 112$	802.816	0
5	Convolution	$128 \times 112 \times 112$	$1.605.632$	73.856
6	Convolution	$128 \times 112 \times 112$	$1.605.632$	147.584
7	Max-pooling	$128 \times 56 \times 56$	401.408	0
8	Convolution	$256 \times 56 \times 56$	802.816	295.168
9	Convolution	$256 \times 56 \times 56$	802.816	590.080
10	Convolution	$256 \times 56 \times 56$	802.816	590.080
11	Convolution	$256 \times 56 \times 56$	802.816	590.080
12	Max-pooling	$256 \times 28 \times 28$	200.704	0
13	Convolution	$512 \times 28 \times 28$	401.408	$1.180.160$
14	Convolution	$512 \times 28 \times 28$	401.408	$2.359.808$
15	Convolution	$512 \times 28 \times 28$	401.408	$2.359.808$
16	Convolution	$512 \times 28 \times 28$	401.408	$2.359.808$
17	Max-pooling	$512 \times 14 \times 14$	100.352	0
18	Convolution	$512 \times 14 \times 14$	100.352	$2.359.808$
19	Convolution	$512 \times 14 \times 14$	100.352	$2.359.808$
20	Convolution	$512 \times 14 \times 14$	100.352	$2.359.808$
21	Max-pooling	$512 \times 7 \times 7$	25.088	0
22	Fully connected	4.096	4.096	$102.764.544$
23	Fully connected	4.096	4.096	$16.781.312$
24	Fully connected	1.000	1.000	$4.097.000$

Table 8.1: Specifications of the VGG16 architecture. The number of learned parameters for the convolutional layers is expressed using (3). The number of learned parameters for a fully connected layer with layer size $H^{(l)}$ and output size $H^{(l)} \times H^{(l+1)}$.

- The total number of learned parameters of VGG19 networks is around 138 millions.

Dropout

- **Dropout** is another popular technique for improving the performance of CNNs. One simple trick for using dropout is to include it after each fully connected layer in the network. This helps to regularize the network and prevent overfitting, and in turn improve the generalization performance on new data.
- Dropout layers are typically deployed after convolutional layers not used after convolutional layers in CNNs, but rather after fully connected layers in convolutional layers, the neurons are typically spatially arranged in a grid, and hopefully each individual neuron should still be able to learn useful structure of the features that follows. However, some researchers have explored alternative forms of spatial dropout that are designed specifically for convolutional layers, so it is an area of ongoing research.

Note:

- It's also important to tune the hyperparameter of dropout, which is the dropout probability, to find the optimal value for your specific network and dataset.
- Additionally, dropout can sometimes be sensitive to the network architecture and the order in which layers are connected, so it's important to experiment with different configurations and combinations to see if performance improves.
- Finally, it's important to note that dropout can sometimes be omitted in small CNNs or shallow networks, as the benefits may not be significant compared to the added computational overhead.

Batch Normalization

- Similar to feed forward networks, batch normalization is a popular technique for improving the performance of CNNs.
- One almighty for using batch normalization is after each convolutional or fully connected layer in the network. This helps to normalize the activations of the network and reduce the internal covariate shift, and thus can lead to faster and more stable training.

Note – Batch normalization can sometimes be sensitive to the batch size used during training, as it's worth experimenting with different batch sizes to see if performance improves.

- It's important to note that batch normalization can sometimes be omitted in small CNNs or shallow networks, as the benefits may not be significant compared to the added computational overhead.

Understanding Inner Layers and Derived Features

- **Visualizing and Understanding Convolutional Networks**" is a paper by Matthew Zeiler and Rob Fergus, which proposes a method for visualizing the feature maps learned by convolutional neural networks.
- The authors start by highlighting the importance of understanding how CNNs work. Specifically, they focus on the visualization of the activation patterns in **the feature maps** produced by each layer of the network. The following figure provides some understanding of the features between the layers. For more details refer to the original paper.

- The network has many channels across multiple layers, and here we present only a few of these channels, focusing on a pair of arbitrary channels within each of the layers 2, 3, 4, and 5.
- Each channel that we visualize has a 7×7 grid of synthesized images (**channel visualization**) as well as a matching 7×7 grid of parts of images from a dataset **original receptive fields**. These channel visualizations and original receptive fields can serve as a visual interpretation of what the specific channel detects.
- The authors also provide a GIF version of the channels visualized in layer 2 (**layer2 simple features** with one channel focusing on edges and another channel focusing on colors.
- As we advance deeper in the network we see that the type of visual patterns detected are much more complex. For example, the two channels presented for layer 4 detect parts of animals, and the channels of layer 5 detect such representations as well.
- Note however, that one of the channels in layer 5 that we present appears to detect either faces or car wheels even though these are very different objects. Hence any attempt to categorize channels based on their "meaning" alone is far from absolute.
- Nevertheless, a visual representation such as this Figure can help to understand the function of individual channels within the network.
- Layer 2 seems to primarily focus on edges and colors, layer 3 detect more complex textures and edges and textures, while later layers more complex features such as object parts and whole objects.

Other Landmark Architectures

- In addition to VGG19 architecture that achieved high accuracy in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014, there are few other landmark architectures. The following lists a small summary of these models.
  - **ResNet** (Residual Networks): ResNet is a CNN architecture that introduced the concept of residual learning, which allows for the training of feature maps from earlier layers by adding skip connections. This enables very deep networks (e.g., 152 layers) to be trained and achieve state-of-the-art performance on image classification tasks.

A shortcut connection (residual connection) as part of a residual network.

  - **Inception** (GoogLeNet): Inception is a CNN architecture that introduced the Inception module, which uses multiple filter sizes in parallel to extract features at different scales. This allows for efficient use of computational resources and achieves high accuracy on image classification tasks.

Inception module

One form of an inception module, playing part in an Inception network. The key idea is parallel computation of various paths followed by a concatenating of the outputs from all paths.

  - The Inception network is a type of convolutional neural network architecture that was introduced in **2014 by Google researchers**.
  - The main idea behind the Inception network is to use multiple filter sizes in parallel at each layer of the network, rather than a single filter size as in traditional CNN architectures.
  - This approach is different from other popular CNN architectures, such as VGG, which typically use one-sized filters in all of their convolutional layers.
  - The Inception network also includes the use of so-called bottleneck layers, which use 1x1 convolutions to reduce the dimensionality of the input before applying larger filters. This helps to reduce the number of parameters in the network, while still allowing it to capture complex features.
  - One of the key advantages of the Inception network is its use of an Inception module, which is a building block for the network that combines the various filter sizes and bottleneck layers. The Inception module allows the network to capture complex and multi-scale features at each layer while also reducing the number of parameters and the computational cost.
  - **MobileNet**: MobileNet is a CNN architecture designed for mobile and embedded devices that uses depthwise separable convolutions to reduce the number of parameters and computations required while maintaining high accuracy. It achieves state-of-the-art performance on mobile, drones and real-time video analysis.
  - **DenseNet**: DenseNet is a CNN architecture that introduces dense connectivity between layers, where each layer receives feature maps from all preceding layers. This promotes feature reuse and reduces the number of parameters required, leading to high accuracy and efficient training.
  - **AlexNet**: AlexNet is a CNN architecture that was the winner of the **ILSVRC 2012**. At the time, it was a breakthrough in image classification, achieving state-of-the-art performance on the ImageNet dataset. It introduced the concept of using **ReLU** activation functions, **dropout regularization**, and **data augmentation** for training deep neural networks. It also used GPU acceleration to speed up training and achieved a significant reduction in error rate on image classification tasks.
  - **Efficientnet**: Efficientnet is a family of convolutional neural network architectures that were developed with the aim of providing better accuracy and efficiency in terms of model size and computation cost. The key idea is to systematically scale up the dimensions of the network's parameters (such as depth, width, and resolution) in a balanced way, while also introducing a new compound scaling method that optimizes these dimensions based on a set of pre-defined constraints.

Performance of various convolutional models as well as efficient net

Beyond Classification

- So far we have looked at how convolutional neural networks with a particular emphasis on the task of **image classification**, e.g. determining if an image is of a cat or a dog.
- However, there are several other important image analysis tasks that are also handled with convolutional neural networks. These tasks deal with analyses and understanding of an image including the location of objects, the count of objects, recognizing between different semantic features of the image, and more.

Note: In terms of the black-box, it is important to keep in mind that not all data is made of monochrome or color images. Within computer vision, one often deals with **image sequences** (short movies), or images that have **more than 3 channels**.

- For example, some images may also have a **distance channel** capturing the distance from the camera per pixel. Further, non-image data can also be handled by convolutional networks.
- Finally, **Graph Neural Networks** (**Geometric Deep Learning** Imagine) also exists in 3-dimensional in nature as it records the state of physical locations (e.g. blood oxygen-level dependent (BOLD) signal in three dimensions over time.

Convolutional Networks and Key Computer Vision Tasks

- As mentioned above, classification tasks are typically deployed at the end of the network. For example, a convolutional neural network $f_\theta(.)$ has output $f = f_\theta(x)$ which is a **vector of probabilities** where the highest probability typically determines the appropriate label for the image.
- As we extend from our detailed study of the VGG16 model and other architectures, **initial layers** of the model $f_\theta(.)$ are typically shared across the different tasks, and the **final layers** are task-specific.
- These final layers **fully** transform the internal latent features of the network into the output vector of probabilities f.
- When one considers tasks other than classification, it is often common to replace the final layers of the network with other layers such that the output is now the desired task.
- The following figure illustrates key computer vision tasks for images.

Figure 4.10: Illustration of some common computer vision tasks: Classification, Classification + Localization, Object Detection, Instance Segmentation. Classification: Assigns a class label to the image. Classification + Localization: Determines the object's location (bounding box) and assigns a class label. Object Detection: Identifies multiple objects in the image, each with a bounding box and a class label. Instance Segmentation: Segments each individual object instance, providing pixel-level masks for each object, distinguishing between different instances of the same class.

- **Object localization** which is the task of identifying the location of an object in an image, as well as possibly the type of object in which case the task is called **localization and classification**.
- **Object detection** which is the task of detecting multiple instances of an object in an image, also separating between the objects and classifying each detected object.
- **Landmark detection**, which is the task of identifying the specific pixel locations of landmarks in an image.
- **Semantic segmentation**: which is the process of classifying each individual pixel to be of a different class in order to have a finer set of classes (pixel wise classification).
- **Instance segmentation**, which finds different instances of objects in the image and separates pixels to be of different instances.
- **Identification** (face recognition), which determines if an image is that of a specific instance (or person).

Let us now consider possible forms of the output f.

- For object localization, it needs to contain information about a **bounding box** which locates the object. This can be in the form of (l_x, l_y, l_w, l_h) where l_x and l_y are the coordinates of (say) the upper left corner of the bounding box and l_w, l_h are the height and width of the bounding box respectively.
- For object detection, a variable number of multiple bounding boxes need to be supported.
- For landmark detection, a list of coordinates of the locations of landmarks comprises the output.
- For semantic segmentation, we need location in the input image, x, has an associated probability vector of classes in the output f. Hence in this case, f can be represented as a tensor with width and height dimensions the same as the input image, and a depth dimension which is the number of classes in the segmentation.
- For instance segmentation, the output is similar to that of semantic segmentation, but instead of recording probabilities of classes, the output tensor records the probability that each pixel belongs to a specific instance of an object.
- Finally, in the case of identification, or face recognition, the task is often cast as a similarity task, where the task is to determine if a face image matches a given pre-stored template or not. The specialty in this case is that the input x is typically compared to **two images**, where one is the template of the person of interest, and the other is a candidate, and the other image is the **current image taken**.

Object Localization

- To get a bounding box object localization assume that we wish to train a CNN that operates on **an input image** x and determines if the image contains a bird or a cat (classes).
- The model's second goal is to determine the **specific location** (l_x, l_y, l_w, l_h) of that object localization. Images with multiple birds or planes are not considered. Images without a bird and without a plane are possible and in this case the output (ratio bounding).

For every to encode the output is

$$f = (\text{bird}_{prob}, \text{plane}_{prob}, l_x, l_y, l_w, l_h)$$

where as in standard classification example $(\text{bird}_{prob}, \text{plane}_{prob})$ is a probability vector, and the other coordinates define a bounding box.

- Here we output the l_{cls} loss, l_{loc} loss, or $l_{cls} + l_{loc}$ loss where l_{cls} measures a prediction of one bird and no plane.
- In terms of training data, for each input image we denote the output as y where image without a bird or a plane is denoted as $y = (0, 0, 0, 0, 0, 0)$ and a bird with a bounding box "do not care" where images with a bird are denoted $y = (1, 0, l_x, l_y, l_w, l_h)$ where the bounding box (l_x, l_y, l_w, l_h) is typically based on a manual determination by a human annotator. Similarly, images with a plane are denoted $y = (0, 1, 0, 0, 0, 0)$ and $y = (1, 1, l_x, l_y, l_w, l_h)$.
- We may consider a loss function such as the squared loss, where $\hat{y}^{(i)}$ and $y^{(i)}$ are the predicted and ground truth for the ith training example, respectively. Hence, $\hat{y}^{(i)}$ and $y^{(i)}$ are both 6-dimensional vectors, and $\hat{y}^{(i)}$ and $y^{(i)}$ are the predicted and ground truth for the ith training example, respectively. Hence, $\hat{y}^{(i)}$ and $y^{(i)}$ are both 6-dimensional vectors, and $\hat{y}^{(i)}$ and $y^{(i)}$ are the predicted and ground truth for the ith training example, respectively.

$$l = \sum_{i=1}^{6} (\hat{y}_i^{(i)} - y_i^{(i)})^2$$

- To perform object localization, say with a model like VGG16, the network can be modified by adding additional layers at the end of the network to output the bounding box coordinates. This can be achieved by adding a fully connected layer with four output units, corresponding to the coordinates of the bounding box. Such simple modifications of networks that were otherwise designed for classification are always possible.

Practice Exercise

Practical 3:

- Open Tutorial 3 of CNN available on the workshop GitHub page. Alternatively, click here.
- Save a copy of this in your Google Colab.
- In this exercise, we build a simple CNN for classification using MNIST dataset.

Practical 4:

- Open Tutorial 4 of CNN available on the workshop GitHub page. Alternatively, click here.
- Save a copy of this in your Google Colab.
- In this exercise, we build a simple CNN for classification using MaxMNIST, a collection of standardized medical image datasets.

Practical 5:

- Open Tutorial 5 of CNN available on the workshop GitHub page. Alternatively, click here.
- Save a copy of this in your Google Colab.
- In this exercise, we use transfer learning strategy for classification using MedMNIST.

Solution Practice Exercise

- Solution Practice 1
- Solution Practice 2