

# Deep Learning

## Recap

Danijel Skočaj

University of Ljubljana

Faculty of Computer and Information Science

Academic year: 2022/23

# Content

---

Basic concepts

Training neural networks

Optimizing training

Convolutional Neural Networks

Computer vision

Recurrent Neural Networks

Transformers

Natural Language processing

Transformers in comp. vision

Generative models

Theory

Applications

SOTA

# Contents

---

**Introduction**: Introduction. Motivation. History. The main concept. Typical solution. Motivating examples.

**Training ANN**: Perceptron. Feedforward neural networks. Loss function. Gradient descend. SGD. Backpropagation. Chain rule. Computational graph. Gradient backward flow. Equations of backpropagation. Backpropagation and SGD. Backprop summary. Backprop example. Quadratic loss function. Cross-entropy loss function. SoftMax. Categorical cross-entropy. Activation functions. Sigmoid. Tanh. ReLU. GELU. Overfitting. Regularisation. L2 and L1 regularisation. Regularised SGD. Dropout. Data augmentation. Pre-processing the data. Weight initialisation. Batch normalisation. Parameter-update optimisers. Momentum. AdaGrad. RMSProp. Adam. Learning rate decay. Learning rate schedules. Setting up the network. Hyper-parameter optimisation.

**CNN**: Convolutional neural networks. Convolution. Convolution layer. Sparse connectivity. Receptive field. Parameter sharing. Translation equivariance. CNNs as FC networks. CNN principle. Stride. Padding. Conv layer parameters. Executing convolution. Pooling layer. Fully connected layer. CNN architecture. Milestone CNN architectures. LeNet-5. AlexNet. VGGNet. GoogleLeNet. Inception. Auxiliary output. ResNet. Residual connection. Ensemble methods. SENet. Feature recalibration. Wide ResNet. ResNext. Parallel pathways. DenseNet. MobileNets. Depthwise separable convolution. Pointwise separable convolution. Xception. ShuffleNet. Group convolution. Neural architecture search. NASNet. EfficientNet. Architecture overview. ConvNext. **Transfer learning**. **Data augmentation**. AutoAugment. CNN regularisation. Cutout. Mixup. Explainability of CNNs. Visualising filters and images. t-SNE visualisation. t-SimCNE. Maximally Activating Patches. Visualising activations. Deconvolution. Guided backpropagation. Occlusion sensitivity. LIME. Class Activation Maps. Grad-CAM. Adversarial images. FGSM and BIM. DeepFool. Accuracy-perturbation curves. Universal perturbations. SparseFool. Adversarial training.

# Contents

---

**Computer vision:** Visual information. Computer vision tasks. Classification. Localization. Semantic segmentation. Fully-convolutional approach. Upsampling. Transpose convolution. FCN. Deconvolution network. SegNet. Encoder-decoder architecture. U-net. Shortcut connections. PSP-Net. Pyramid Pooling. DeepLab. Atrous convolution. Atrous Spatial Pyramid Pooling. Semantic segmentation overview. Beyond segmentation. Surface-defect detection. Counting. Obstacle detection. Image enhancement. Spatially Adaptive Filter Units. Semantic edge detection. Object detection. Detection of multiple objects. Two-stage object detectors. Region proposal methods. R-CNN. Fast R-CNN. RoI pooling. Faster RCNN. Region Proposal Network. Instance segmentation. Mask R-CNN. RoI Align. Task specific heads. Human pose detection. Panoptic segmentation. Feature Pyramid Network. Panoptic Feature Pyramid Networks. Single-stage object detectors. SSD. Anchors. Yolo. RetinaNet. Focal loss. FCOS. Traffic sign detection. Face detection. Object detection architectures overview. Performance of object detectors. Classification performance measures. Object detection performance measures. Semantic segmentation performance measures.

**Recurrent Neural Networks:** Sequential data. CNN-based approach. Requirements. Recurrent Neural network. Unfolding RNN. Multilayer RNN. Recurrence formula. Computational graph. Backpropagation through time. TBPTT. Character-level language model. Predicting the next letter. Interpreting the hidden cell values. RNN trade-offs. BPTT problems. Long Short Term Memory. LSTM gradient flow. Gated Recurrent Units. RNN variants. Bidirectional LSTM. LSTM examples. Attention mechanism. Attention in RNNs. Image captioning with attention.

**Transformers and NLP:** Transformer architecture. Encoder. Self-attention. Multi-head attention. Positional encoding. Decoder. Encoder-decoder attention. Masked self-attention. Decoding. RNNs vs. Transformers. Transformer-XL. BERT. BERT pretraining. BERT fine-tuning. RoBERTa. ALBERT. DeBERTa. BERT examples. T5. GPT. Unsupervised generative pretraining. Supervised discriminative fine-tuning. GPT-2. GPT-3. InstructGPT. Alignment problem. ChatGPT. Power and considerations of modern LLMs. Speech recognition. Whisper. Music transformer.

# Contents

---

**Transformers in computer vision:** Image transformer. ViT Vision transformer. DeiT. MViT. Swin transformer. Swin V2. SeMask. CvT. CoAtNet. DETR. Deformable DETR. UP-DETR. MaskFormer. Mask2Former. DINO. DINOv2. MLP-Mixer. Segment Anything.

**Generative models:** Gaussian and GMM. Principal Component Analysis. Robust projection. Autoencoders. Masked autoencoders. Autoencoders for anomaly detection. Generative vs. discriminative approaches. Variational Autoencoders. VAE generating images. VAE in music. Generative Adversarial Networks. Training GAN. DCGAN. LSGAN. Wasserstein GAN. Improved Wasserstein GAN. Progressive GAN. Pix2pix cGAN. CycleGAN. BigGAN. SinGan. StyleGAN. F-AnoGAN. GANomaly. C-VTON. Normalizing flows. Real NVP. Flow++. Glow. DifferNet. FastFlow. PixelCNN. VQ-VAE. CLIP. DALL-E. Diffusion models. DDPM. Stable Diffusion. GLIDE. DALL-E2. ControlNet. Midjourney.

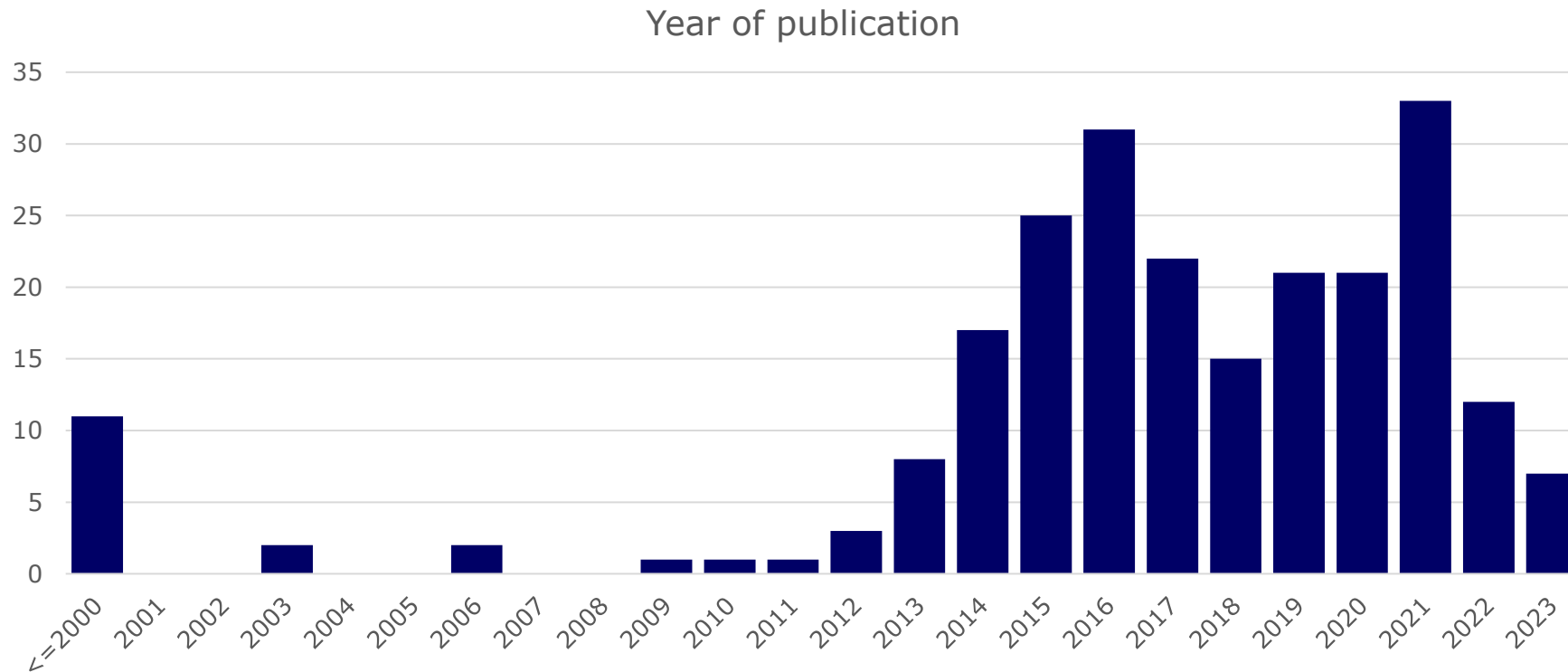
**Conclusion:** Deep learning. Key factors for fair deep learning. Learning regimes. Deep reinforcement learning. RL for local navigation. Built in vs. Learned. Beyond CV and NLP. Function approximator. Development, deployment and maintenance. Problem solving.

Level of details:

- Detailed knowledge
- Main concepts
- Main idea

# Research papers

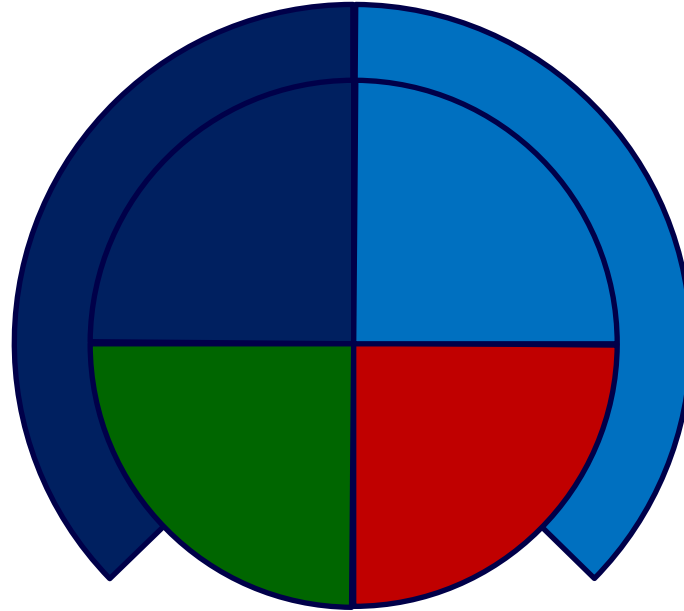
- 233 papers presented/mentioned
- All of them available online!
- Seminal papers and SOTA



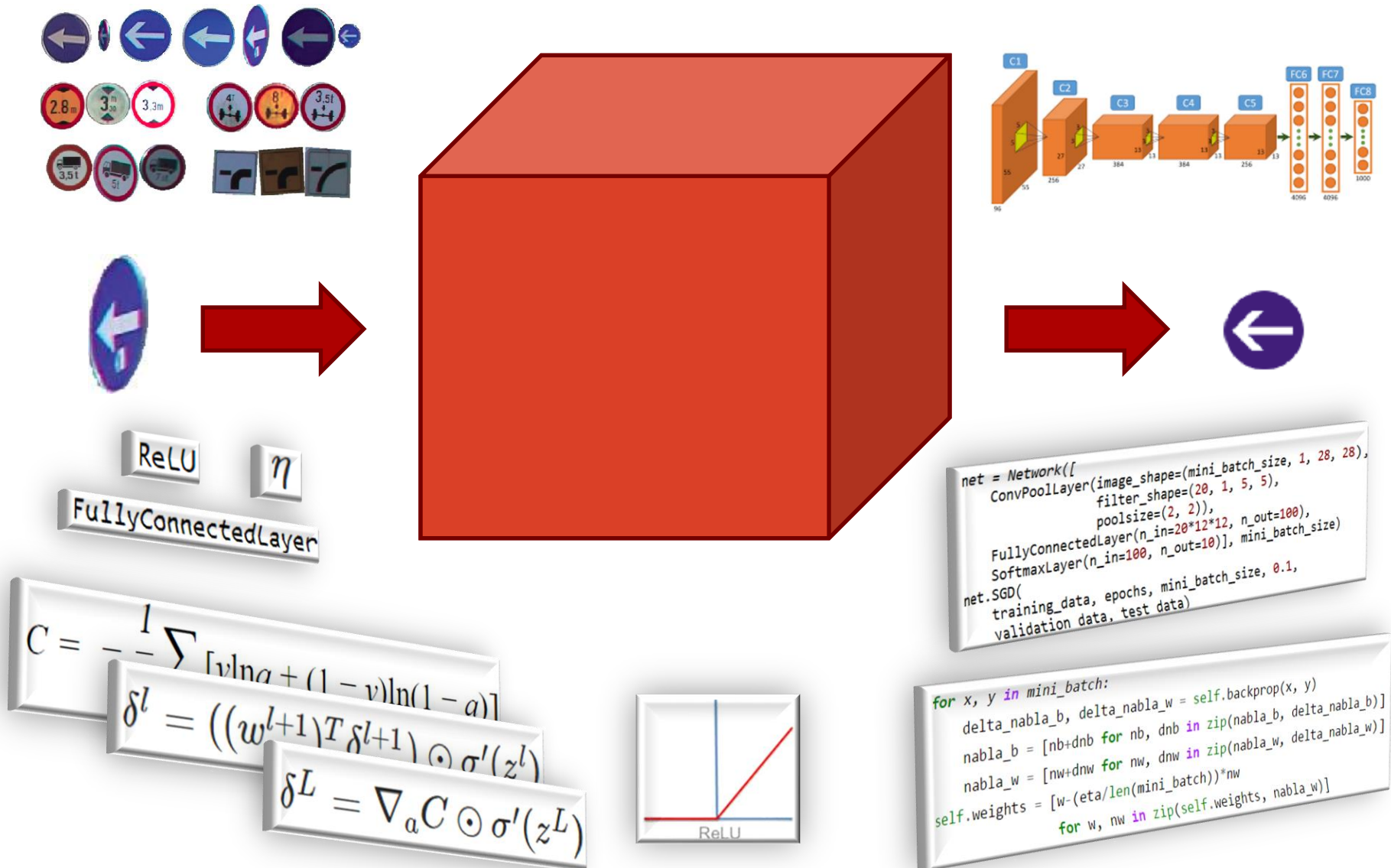
# Grading

---

- Assignments: 25 pts
  - Project: 25 pts
  - Written exam: 25 pts
  - Oral exam: 25 pts
- 
- The student has to collect at least 50% in every task

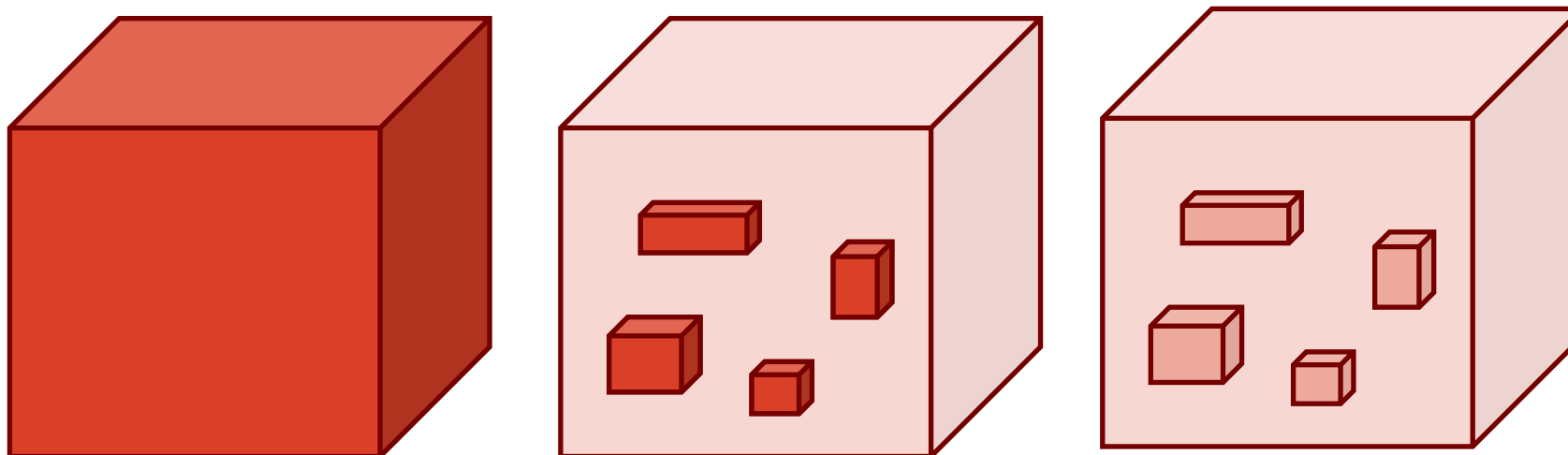


# Understanding deep learning





# Using deep learning



Spread of usage

Problem complexity

Course goal