week01

CV in self-driving cars, face recognition, GAN

SNN（脉冲神经网络）

CNN

1. Data：Augmentation, Selection Strategy
2. Network: Initialization strategy, learning strategy, optimization strategy, framework-modeules, Layers.
3. Loss: Loss type, loss selection, update strategy, evalution

CV

1. low level: 从图像能直接得到的信息，basic operations, Convolution
2. mid level：链接high level和Low level的桥梁，feature extraction, Feature points, Corner point, SIFT
3. high level：bounding box，从2维怎么构建3维信息，real CV tasks by using CNN, classification

Fantastic Transformation, Style Transfer, Image Generation, 2D To Depth, 3D objects from 2D, Slam

pandapow

Research: AutoML, Acceleration,

Applications: **Image**/Video Classification, Object Detection, Keypoint Detection, **Recognition**, **Segmentation**, Voxcel, Tracking, 2D-3D/3D-2D, Image Captioning, **Image Transfer**, Mixed Inputs, GAN, Slam.

Engineering: Modified Models, **Light Models,** Accelerating Algorithm

Python主要针对有research意愿的人。**C++**主要针对有工程意愿的人。

Caffe、**PyTorch**上手快

**Coding Preparation**: 200/300+题，medium难度40min做对4/5道，Lintcode/Leetcode

LintCode：<https://www.lintcode.com/problem/> 账号：wjq332608421 密码：wjq613613

Tiny Project, Kaggle：Cancer Detection, Cactus Identification, Protein Atlas Image Classification.

Reflection Removal, Super Resolution, Face Frontalization, Add/Remove Sth

Video -> Sound, Video + Sound -> Classification, 2-Stream Videos -> Classification

CVPR, ECCV, ICCV – 偏应用

NIPS – 偏理论

PAMI – 顶刊，较滞后

SIGGRAPH, arXiv, Kaggle

冈萨雷斯-《数字图像处理-第三版》

合恩-《矩阵分析》

Bishop-《Pattern Recognition and Machine Learning》

week02

Convolution:

1-dim convolution: kernel (需180度翻转)，相关性操作correlation (不需180度翻转)：1）kernel很多时候是对称的，翻不翻转都是一样的；2）操作行为类似；3）kernel的值是学习学出来的，对转不转是没有意义的。

2-dim convolution: 补0（padding）数为kernel/2取整的值时，卷积后的数据大小不变。

Image convolution: 图像分不同的channel，每个channel都是一个2维的，每个channel需分别和各自的kernel做卷积。

First-order derivation（一阶导）：相邻像素的差。x为坐标，f(x)是灰度值。相邻 – 人为规定，4邻或者8邻。

Second-order derivation（二阶导）

上式的严格定义如下：

– 拉普拉斯算子

一阶导、二阶导在图像中找信息，能找到角点、边缘。

边缘（Sobel）：二阶导更敏感，二阶导能产生双边缘的效果。-> Sharpen（锐化，Laplacian）：让图像看起来更加清晰（=原图+边缘）。

卷积能快速算一阶导、二阶导，所以用卷积的方法。

钝化：让图像更加模糊。 -> Gaussian Kernel（加权平均，平滑处理）

一阶导：

二阶导：

二阶导拆分成两次一阶导，计算速度更快。

Bilateral Filtering：非边缘部分变平，同时保留边缘。

Feature Points/descriptor：represents pf objects/pixels.

什么算好的特征点？

1）非常具有信息量；

定义一个window：w

指窗口的整体平移量。V的值就是平移后和平移后的灰度值相减。

如果(x,y)对应的是角点，那么新的窗口对于(x,y)就有更大的响应。

泰勒展开：

为高阶无穷小项。

若上式大于某一个阈值thre，则判断其为角点。k的值一般为0.04。为矩阵的特征值。其中：

2）旋转不变性、对光亮不敏感

3）尺度不变性（距离远近看一个物体，物体类别不会变）

SIFT：Scale-Invariant Feature Transform.

1. Generate Scale-Space: DoG（Difference of Gaussian）：用一系列的高斯核对图像进行平滑，模拟出不同尺度下图像的样本，然后将模拟出的高斯图像做差，得到所谓的轮廓线；在不同尺度的空间下找到极值点，找到极值点后再对极值点进行精确的定位，消除所谓的边缘棱角，然后生成descriptor。-> 在尺度层面上找边缘。
2. Scale-Space Extreme Detection：尺度空间内的极值点（该点跟周围的点相比，最大或最小）。20m看到的图中的点，除了与自身所在图相邻的点进行比较外，还要与10m、40m看到的图像对应的点进行比较。
3. Accurate Keypoint Localization



指尺度。

1. Eliminating Edge Responses：海森矩阵。

DoG：too sensitive to edge, unstable.

1. Orientation Assignment：解决旋转不变性。
2. Keypoint Descriptor：对角度、尺度、边缘没有影响，而且非常精准。

FAST（faster version of sift）、SURF、BRIEF、HoG（detection）、OrB（SLAM）

RANSAC

homegraphy matrix

Feature Pyramid Networks：尺度上的变化有金字塔的效果。

SIFT的流程：如何做到尺度不变、旋转不变的。

泰勒展开

插值问题

Bag of Words（BoW）

Image Stitching

1. Find feature points in each image
2. Use RANSAC to find keypoint matches
3. Use homography matrix to get transferring info
4. Merge two images.

week03

well-posed（1、A solution exists; 2、Solution is unique；3、Solution’s behavior is continuous） \ ill-posed（如给黑白照片上颜色，解不唯一）

Supervised Learning

1、Who supervises who? A standard supervises the system.

Regression: Object Detection, Keypoint Detection（人脸关键点检测）, Prediction。

Classification: Image/Video Classification, Segmentation（分割，Instance Segmentation, Semantic Segmentation）。

识别 \ 分类 \ 检测（Bounding Box）

Unspuervised Learning（没有标准答案）

Q: How to type a video within 1,000,000,000 videos in a short time?

A: Classification. 1. 标注成本问题。

A: Clustering: 生成树状结构

**Linear Regression**：训练的过程就是寻找参数的过程。

Cost：hypotheses和reality之间的差距，目的就是最小化Cost。

Gradient Desceng Algo

while not converge {

}

过大，容易导致梯度爆炸。

Extension – Multiple Variables：

表示第i个样本。

Feature Scaling

**Logistic Regression**: Classification

Cost Function（二分类Cross entropy，对于多分类使用Softmax）:

Sigmoid函数：，其导数为：

Multi-Classes: Softmax

week04

**Neural Network**

Activation Function：非线性作用。

：activation of unit in layer 。

神经元个数怎么确定？经验。

Classification：其输出表示属于哪个类别的概率。

：sample ID；

：# of samples;

：output ID

：# of outputs

若层数较多，通过Gradient Descent求参数变得不现实了。

**Back Propagation**：Chain Rule。（会推导）

Chain Rule存在的问题：

1. 梯度消失/爆炸。

**Regularization（惩罚项）**：防止过拟合，相当于让值变小，让拟合曲线变得更加平滑。

对于线性回归，加入惩罚项后，Gradient Descent项变为：

L2 Regularization：，Ridge Regression

L1 Regularization：，Lasso Regression

L2和L1的异同、特点？

|  |  |
| --- | --- |
| L2 regularization | L1 regularization |
| Computational efficient due to having analytical solutions | Computational inefficient on non-sparse cases |
| Non-sparse outputs | Sparse outputs |
| No feature selection | Built-in feature selection（能使一些参数为0） |

week05

RANSAC算法用于拟合和匹配：<https://www.jianshu.com/p/4b5de31a8894>

fast median filter

图像分割：阈值法（cv2.threshold）、漫水法（cv2.floodFill）、水域法、金字塔分割、均值漂移分割、背景减除。

阈值法：大津阈值（cv2.THRESH\_OTSU）、自适应阈值。

图像滤波：均值滤波、加和滤波、中值滤波、高斯滤波、双边滤波、盒子滤波。

边缘检测：Sobel、Scharr、拉普拉斯、Canny、形态学梯度。

遥感图像：GDAL库。

医学图像：呈层级数据

彩色空间：RGB、CMY、CMYK、HIS。

图像变换：拉伸、缩小、Warp、旋转；一致调整变换或均匀调整变换、图像金字塔、非均匀映射、仿射变换、透视变换；极坐标映射、对数极坐标、任意映射。

图像常见的传统特征描述：LBP、Gabor、HOG、SIFT、Hough变换，颜色直方图、频谱特征提取、圆周频率滤波，矩特征、Harris角点检测

week06：SVM、非监督学习

**SVM（Support Vector Machine）**

，

，，系数不影响平面性质，可将替换为1。截距项对应。

上两式可转换为：

两个平面之间的margin为：

等价于：，且需满足。

拉格朗日乘子法：将约束条件与目标函数融合在一起。

Soft margin：，相当于将平面整体平移了一下。

其中，越大，对误分类的惩罚越大，更可能过拟合。其可看作正则化系数的倒数关系。

Kernel：能产生非线性的效果，类似激活函数。

令，其为高斯核。

越大，kernel更平滑，拟合能力更低，bias会增加，variance会减小。

只要样本越趋近于，值越趋近于1；越远离，值越趋近于0。

linear、polynomial、Gaussian（RBF）、Laplacian（在高斯核基础上去掉2次方）、Sigmoid。

n：# of features；m：# of samples.

Linear Regression

Logistic Regression / Classification

L2 Loss

Sigmoid Function

Cross Etropy / Multi-Label

Neural Network

L1 / L2 Regularization

Overfit / Zigzag

BP / Cost Function / Hypothesis

Gradient Vanishing ./ Explosion

SVM / Kernel / Derivation

Normal Equation

Unsupervised Learning

**K-Means**

收敛：新找德中点与之前德中点很靠近。

需要自己手动实现。

初始化方法：K-Means++[2007, Arthur & Vassilvitskii]

离得远，被选为新中心点的概率大。

K chosen：elbow method / ISODATA

参考sklearn各种k-means方法

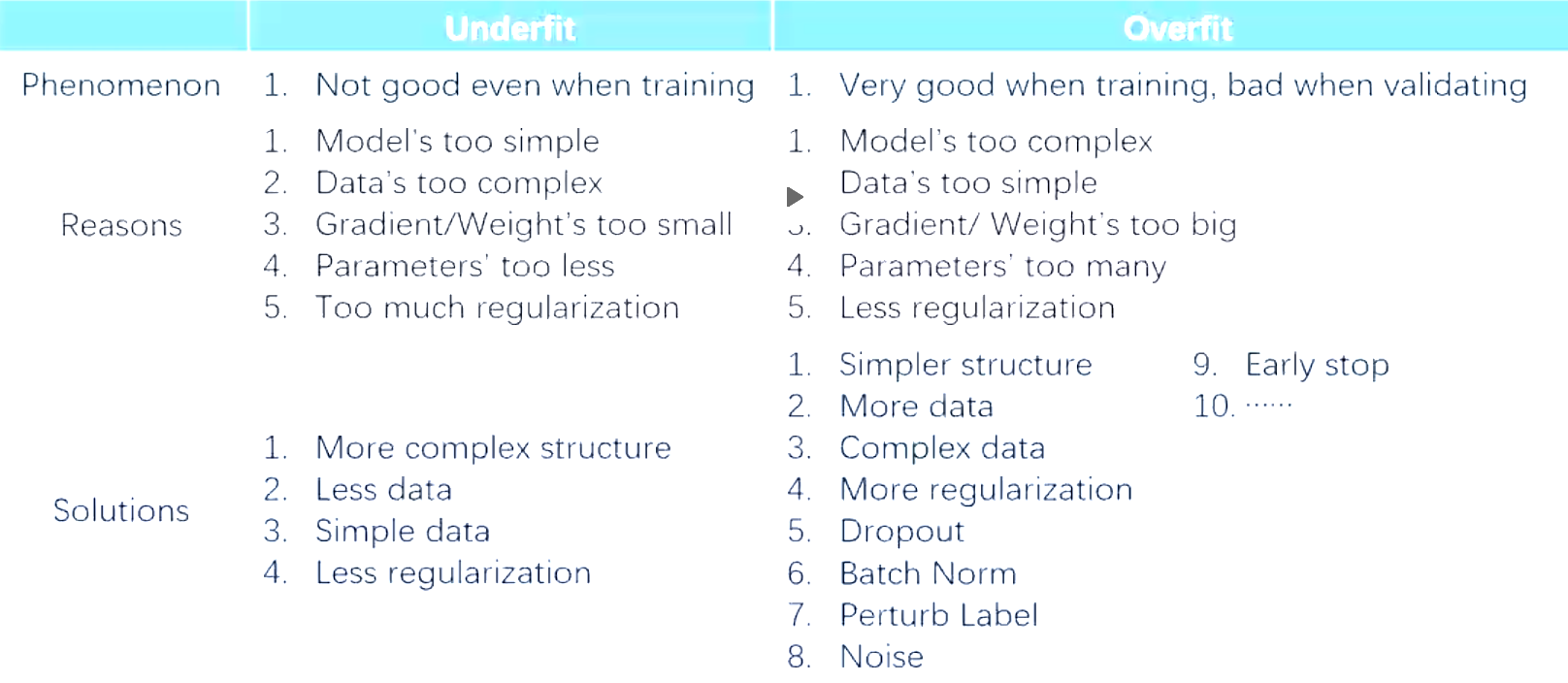
Training / Validation / Test Set

Validation：1）训练时起到测试的作用；2）训练超参数。

交叉验证：将训练集分成9份，拿出一份当作Validation，这样验证9次。以此来决定超参数。将此方法训练得到的模型用于Test数据上。

**Underfit / Overfit**

有时候训练一定轮次之后，training loss先减小后增加，亦出现Overfit。



Perturb Label：按样本0.1比例将实际的标签为1，将其标为0，即进行随机标注。

High Bias: Underfit.

High Variance: Overfit.

正则项的越小，越容易过拟合，产生的Loss越小。

|  |  |
| --- | --- |
| To reduce high variance | + |
|  | -complexity |
|  | +samples |
| To reduce high bias | - |
|  | +complexity |

Decision Tree: ID3 / C4.5 / CART（Gini Index）

week07 Outline & Layers

分类、**检测**、Image Transfer。

CV两个项目选一个，具体项目是？

1. 图像拼接；2）LocNet；3）动物多分类实战；4）3D BBX；5）人脸关键点检测；6）人脸转正；7）照片反光移除。

convolution + pooling layers + 非线性

27:00