RECENT TRENDS IN ARTIFICIAL INTELLIGENCE BASED COMPUTER VISION

Prof. Dr.-Ing. Achim Ibenthal

Hochschule für angewandte Wissenschaft und Kunst
Fakultät für Naturwissenschaften und Technik, Göttingen

SoSe 2024

Ziel: Understanding and applying AI models.

Inhaltsverzeichnis

Contents of the presentation I

1. Zero-shot learning.
2. One-shot learning.
3. Few-shot learning.

Lecture 1

Prof. Dr.-Ing. Achim Ibenthal

Hochschule für angewandte Wissenschaft und Kunst
Fakultät für Naturwissenschaften und Technik, Göttingen

SoSe 2024

Ziel: Verständnis Lecture 1

Nutzungsrichtlinie

Die Inhalte dieser Vorlesung entstammen hauptsďż˝chlich den Quellen [1], [2], [3], die als vorlesungsbegleitendes Material empfohlen werden. Weitere Quellen sind im Text kenntlich gemacht. Darďż˝ber hinaus gehende Inhalte und Bilder sind © A. Ibenthal.

Die Nutzung dieses Dokumentes dient ausschlieďż˝lich der Lehre und ist den Studierenden der HAWK, Fachhochschule Hildesheim, Holzminden, Gďż˝ttingen fďż˝r die Dauer ihrer Immatrikulation gestattet. Die Weitergabe, Verďż˝ffentlichung oder kommerzielle Nutzung – auch nur auszugsweise – ist nicht gestattet.

Hinweise zu Fehlern, fehlenden oder falschen Referenzen, sowie weitere Kommentare und Anregungen sind stets willkommen und werden zur Verbesserung des Vorlesungsmaterials fortlaufend eingearbeitet.

Übersicht

zero-shot learning

Definition: Zero-shot learning.

aims to train a model that can classify objects of unseen classe(target domain) via transfering knowledge obteined from other seen classes(source domain) with the help of semantic information

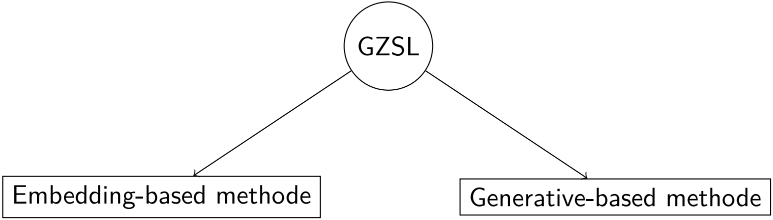
Problem

more challenging scenario with no data for unseen classe,relying solely on descriptions and transfered knowledge from seen classes.

Definition: Generalized Zero-shot learning.

more realistic scenario with some data for seen classes to help classify unseen classes using description and visual representation

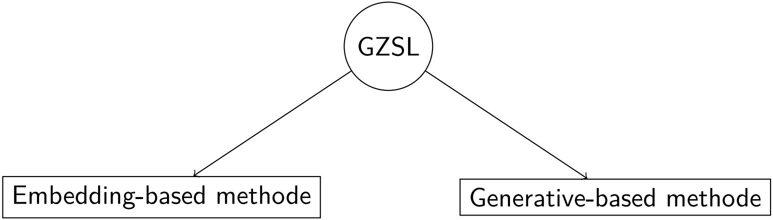
Generalized Zero-Shot Learning (GZSL)



Definition: Embedding-based methode.

learn an embedding space whether visual-semantic ,semantic-visual, commun/latent or a combination of them (Graph-Based, Autoencoder-Based, Meta-Learning-Based, Compositional Learning and Bidirectional Learning-Based)

Generalized Zero-Shot Learning (GZSL)



Definition: Generative-based methode.

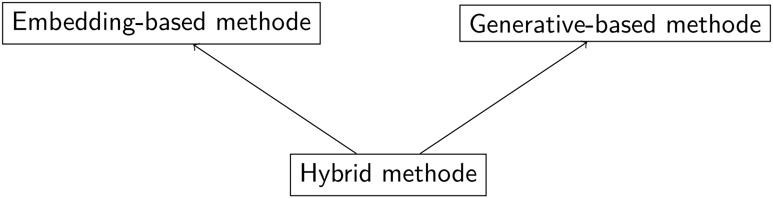
convert GZSL into a conventional supervised learning problem by generating visual features for the unseen classes (Generative Adverserial Network and Variational AutoEncoder (VAE))

Generalized Zero-Shot Learning (GZSL)

	Embedding-based	**Generative-based**
Adventage	• can be effective for simple objects	• can potentially handle complex object better by creating realistic visual representation
Desadvantage	• suffer from the seen classes overfitting problem due to the data imbalance nature of ZSL	• the quality of generated images can be a chalange and impact classification accuracy

Tabelle: Advantage and Desadvantage of GZSL methodes

Generalized Zero-Shot Learning (GZSL)



- The Constractive Embedding for generalized zero-Shot learning paper proposes a hybrid GZSL framwork that comnbines the two models

Definition: Constractive Embedding.

is integrated into hybrid GZSL with the goal to construct embeddings that facilitate the recognition of unseen classes by leveraging the relationships and similarities between seen and unseen classes

Generalized Zero-Shot Learning (GZSL)

Definition: Constractive Embedding.

is integrated into hybrid GZSL with the goal to construct embeddings that facilitate the recognition of unseen classes by leveraging the relationships and similarities between seen and unseen classes

Constractive Embedding for GZSL

Class-wise Supervision

instance-wise Supervision

Definition: Class-wise and instance-wise supervision .

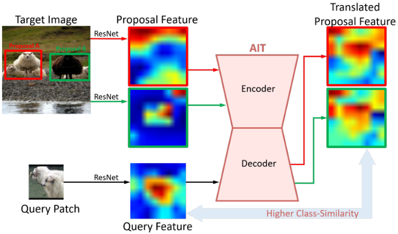
Class-wise Supervision is to classify the membership and the instance-wise Supervision is to identify the similarities or the differenz between individual image .

Übersicht

One-shot learning

Definition: One-shot learning .

aims to identifying within a target image all object instance of the same class,implied by a query image patch with the problem that its label and its respective exemple are not available in the training data .



Definition: Adaptive Image Transformer .

AIT module aims to compare the features of regions or abjects in the Target Image based on the language translation process:

One-shot learning

- AIT can adaptively represent each region proposal so that the similarity with the query can be evaluated ,that module uses an attention based encoder-decoder architecture to simultaneausly explore intra-coder and inter-coder
- encoder converts the inpute data (e.g. features of objects proposals) into internal representation.
- decoder generates the output from this internal representation(e.g. class similarity between proposals and the query image.)
- Intra-coder attention refers to the relationships within a single coder, for exemple the relationship and the importance between different parts of a single image proposal to generate a better internal representation.
- Inter-coder attention examines the relationshipes between different coders. the model learns how each proposal is related the query image,which help to better judge class similarity

One-shot learning

- sparse attention mecanism help the model focus on the most important part of the data.

Evaluation Higher mAP (Mean Average Precision) means more accurately detecting and classifying abjects from all classes .

One-shot learning in Text-to-video generation

Generation of videos from text

Traditional approaches

extensive and large Text-video datasets are requires to train a model

One-Shot Video Tuning

learns to generate videos based on just one exemple instruction and its corresponding video

One-shot learning in Text-to-video generation

One-shot video tuning

`" a man is running on the beach" and a single video that shows exactly this scene .`

1. pretraining phase : learns to generate images from text description
2. learn how to translate the specific text description into motion pictures, das model learns that the beach describe a specific environment and a man has a certain shape
3. generating similar video : generation of videos based on different text description with different scenario , event though had learnt just from one exemple

Application

1. object editing .
2. Background changing .
3. generating similar video : generation of videos based on different text description with style transfer for exemple from real-worls into comic style.

Übersicht

Few-shot learning

Definition: Few-shot learning .

aims to learn information about Data from very limited number of training data when the collecting a large dataset is impossible .

Problem

more challenging for few-shot learning in traditional neural network is the low effectiveness with very limited labled data (e.g. overfitting).

Definition: Overfitting .

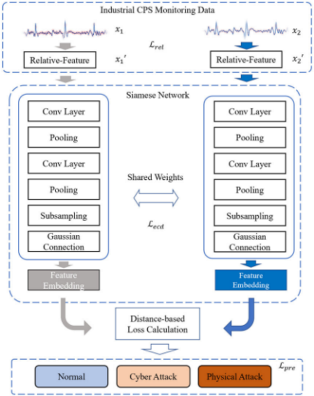
A model overfits the training data when it describes features that arise from noise or variance in the data, rather than the underlying distribution from which the data were drawn. Overfitting usually leads to loss of accuracy on out-of-sample data.

Few-shot learning

Definition: A few-shot learning model based on Siamese Convolutional Neural Network (CNN) .

A Siamese CNN consists of two identical subnetworks that work in parallel and share the same weights. This network compares two inputs and learns whether they belong to the same class or not. By comparing sample pairs, the model can learn more robust and generalizable features, which reduces overfitting.

Few-shot learning



- pooling layer helps in reducing high dimensional convolutional features .
- Conv layal to identify from the imput data the importent features using edges and textures detection .
- two combinations of convolution layer and pooling layer are introduced to extract feature embeddings

Few-shot learning

- the distance between these two feature embeddings will be calculated to identify whether these two input samples belong to the same class.

Few-shot learning

Definition: Few-Shot Hyperspectral Image Classification .

Hyperspectral Image Classification : The objects are classified by its HyperSpectrum .

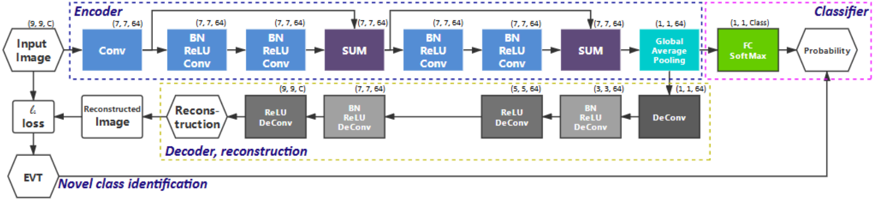
Problem

Current hyperspectral image classification assumes that a predefined classification system is closed and complete, and there are no unknown or novel classes in the unseen data. When a new class appears, it will be interpreted as an error ,what is not the case .

Definition: MDL4OW(multitask deep learning).

method that simultaneously conducts classification and reconstruction in the open world .The reconstructed data are compared with the original data; those failing to be reconstructed are considered unknown classes .

Few-shot learning



- the encoder/features extractor get from the inpute image the proposal features using Pooling and convolution layal .
- after extracting the latent features the function softmax serves as the classifier and outputes the probability to the known classes .
- the reconstruction task uses the deconvolutional layers to increase the spatial dimonsion of the latent features gradually, the output should be similar to the input data .

Application : Generative-based methode(WGAN) I

- The application is coming to solve the problem that we dont have enough data from the unseen classes so the generative-based mothod is coming to generate more data .

Objective The ability of the model to generate new feature vectors that represent unseen classes allows to extend classification to unseen classes.

Application : Generative-based methode(WGAN) I

Definition: Generative space.

By training a generative adverserial network(GAN), they learn to generate new features vectors.

- Our application is about a simplified [AWA2] databases instead of using raw images, extracting features that present the image in a compact yet informative form.

AWA2 contains 30475 images cotegorized into 40 seen classe for training and 10 additional unseen classes for test.

Attribute AH-splits.mat contains the attribute data,every class is represented by a vector with 85 attributes

- The values can be benary or float, which can represent the strength of these properties

Application : Generative-based methode(WGAN) II

Features res101.mat representes the processed image characteristics for each image in AWA2 database, res101 has the ability to present images in a very compact and information-rich manner, rather than raw images that are data-wise heavy.the features are stored as vectors: each vector represents an image and contains a numerical description of the image e.g. shapes, colors, textures and structures .

Application : Generative-based methode(WGAN) I

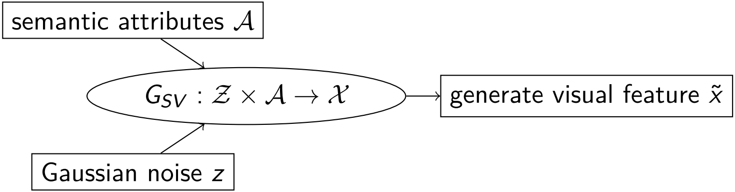
Problem

Let $S = \{(x_i^s, a_i^s, y_i^s)_{i=1}^{N_s} | x_i^s \in X^s, a_i^s \in A^s, y_i^s \in Y^s\}$ and $U = \{(x_j^u, a_j^u, y_j^u)_{j=1}^{N_u} | x_j^u \in X^s, a_j^u \in A^s, y_j^u \in Y^s\}$ represent the seen and unseen class data sets, respectively, where $x_i^s, x_j^u \in \mathbb{R}^D$ indicate the D-dimensional images (visual features) in the feature space X^s that can be obtained using a pre-trained deep learning model such as ResNet. a_i^s,$a_j^u \in \mathbb{R}^K$ indicate the K-dimensional semantic representations (attributes vector) in the semantic space A. $Y_s = \{y_1^s, ..., y_{40}^s\}$ and $Y_u = \{y_1^u, ..., y_{10}^u\}$ indicate the label sets of both seen and unseen classes in the label space Y where 40 and 10 are the number of seen and unseen classes, $\mathcal{Y} = Y^s \cap Y^s$ denotes the union of both seen and unseen classes and $\emptyset = Y^s \cup Y^s$

Application : Generative-based methode(WGAN) I

Problem

GANs generate new data samples by computing the joint distribution $p(y, x)$ of samples utilizing the class conditional density $p(x \mid y)$ and class prior probability $p(y)$.GANs consist of a generator $G_{SV} : \mathcal{Z} \times \mathcal{A} \to \mathcal{X}$ that uses semantic attributes $\mathcal{A}$ and Gaussian noise $z \in \mathcal{Z}$, to generate visual feature $\tilde{x} \in \mathcal{X}$, and a discriminator $D_v : \mathcal{X} \times \mathcal{A} \to [0, 1]$



Application : Generative-based methode(WGAN) I

the discriminator

the discriminator $D_v : \mathcal{X} \times \mathcal{A} \rightarrow [0, 1]$ distinguishes real visual features x from the generated ones $\tilde{x}$ and return a value between 0 and 1, which indicates whether the features are real or generated.

the discriminator

When a genera tor learns to synthesize data samples for the seen classes conditioned on their semantic representations $\mathcal{A}_s$, it can be used to generate data samples for the unseen classes through their semantic representations $\mathcal{A}_u$. However, the original GAN models are difficult to train, and there is a lack of variety in the generated samples.In addition, the mode collapse is a common issue in GANs, as there areno explicit constraints in the learning objective.

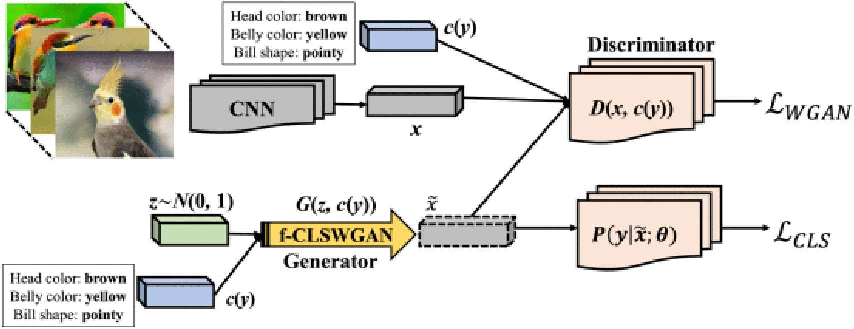
Application : Generative-based methode(WGAN) II

Definition: mode collaps.

the generator produces only very limited variation of data ,that not enough for the classification of varied animales. to ensure that generated features are useful for the classification.

- To over come this issue and stabilize the training procedure, many GAN models with alternative objective functions have been developed

Application : Generative-based methode(WGAN) I



- It provides a measure of the discrepancy between predicted probabilities and true labels, and serves as an objective function to guide model training. By minimizing the negative log-likelihood loss, the model learns to make more accurate and confident predictions.

Satz \nsatz and Theorem \ntheorem I

Definition: wasserstein-distance and gradient penality.

$$\mathcal{L}_{WGAN} = E[D(x^s, a^s)] - E[D(\tilde{x}^s, a^s)] \\ - \lambda E[(\| \bigtriangledown_{\hat{x}} D(\hat{x}, a^s)\|_2 - 1)^2],$$

the difference between the expectation that the descriminator classifies real data as real and generated data as generated minus the gradient penality times the power λ .

- In addition, to generate discriminative features, the negative log-likelihood is used to minimize the classification loss

Satz: negative Log-likelihood.

$$\mathcal{L}_{CLS} = -E_{\tilde{x}^s \sim p_{\tilde{x}^s}}[\log P(y^s | \tilde{x}^s; \theta)], \tag{4}$$

Application : Generative-based methode(WGAN) I

Definition: wasserstein GAN (WGAN).

instead using the standard GAN Lossfunction, where the generator tries to deceive the descriminator, the wasserstein distance is used to stabilize the training and make it converges faster.

Definition: gradient penality.

ensure that the gradient of the descriminator on $\tilde{x}$(mix of real and generated data) satisfied the lipschitz condition , this penality is necessary to stabilize the convergence of the WGAN .

Satz: Lipschitz condition.

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ satisfies the Lipschitz condition if there exists a constant $K \geq 0$ such that for all $x_1, x_2 \in \mathbb{R}^n$:

$$\|f(x_1) - f(x_2)\| \leq K\|x_1 - x_2\|$$

Satz \nsatz and Theorem \ntheorem I

Satz: negative Log-likelihood.

The final objective function can be written as:

$$\min_G \max_D \mathcal{L}_{WGAN} + \beta \mathcal{L}_{CLS}, \tag{5}$$

- The $\min_G$: pushes the generator to minimize the WGANloss by generating a fake simple that resemble real data .
  $\max_D$: pushes the descriminator to maximize the WGANloss by excluding the maximum of fake simple .

Application: MDL4OW (Multitask Deep Learning) I

Problem

In a closed-world setting, the existence of unknown or novel classes will lead to false positives, thereby reducing the precision of a model. The classifier tries to classify the unknown classes as seen classes because it cannot differentiate between known and unknown classes.

Problem

Let X be the sample instance space. Each instance $x \in X$ is given a limited training set with index k, consisting of only a few samples (x^k, l^k), where $l^k \in L = \{1, \ldots, |C|\}$ is the label index for x^k.
The multilevel convolutional layers, along with batch normalization and the Rectified Linear Unit (ReLU), serve as a good encoder $\phi(\cdot)$ to extract the spectral-spatial features x_ϕ as the representation of the sample instance:

$$\boxed{x_{\phi_{conv}}} \longrightarrow \phi(\cdot) \longrightarrow \boxed{x_{\phi_{conv}} = \phi_{conv}(x)}$$

Application: MDL4OW (Multitask Deep Learning) II

Definition: Normalized Inputs

Normalization adjusts the values so they have a mean of 0. This reduces fluctuations in the data (e.g., from $(-100, 500)$ to $(-1, 1)$), making the training process easier.

Problem

Then, the classifier $f(\cdot)$ takes the output vector x_ϕ from the feature extractor $\phi(\cdot)$ as its input.
In a pure deep learning scenario, the fully connected layer with the SoftMax activation function serves as the classifier $f(\cdot)$ and gives the probability $P(y = j|x_\phi)$ of the j-th category:

$$P(y = j|x_\phi) = \frac{\exp\left(x_\phi^T w_j + b_j\right)}{\sum_{c=1}^{C} \exp\left(x_\phi^T w_c + b_c\right)}, \tag{1}$$

Application: MDL4OW (Multitask Deep Learning) I

- w_j is the weight vector of the j-th neuron in the fully connected layer
- b_j is a bias element corresponding to the j-th neural
- C is the number of the category
- The classification task is to find the optimal parameters for the network by minimizing the cross-entropy loss function ℓ_c:

$$\ell_c\left(y, \hat{y}\right) = -\sum_{i=1}^{C} y_i log\left(\hat{y}_i\right), \tag{2}$$

- C is the number of predefined classes
- y is the ground truth label, and $\hat{y}$ is the predicted label.

Application: MDL4OW (Multitask Deep Learning) I

- a naive solution to identify unknown classes is considering those instance with the largest probability smaller than 0.5 as unknown (softmax function with threshold$= 0.5$)
- the network cannot still identify the unknown. To empower the network with this ability, we add a reconstruction task:

$$\hat{x} = f_r\left(x_\phi\right), \tag{3}$$

- $\hat{x}$: the reconstructed instance
- $f_r(\cdot)$ the reconstruction function or named the decoder.
- x_ϕ is the output latent features from the encoder $\phi(\cdot)$.
- the distance ℓ_1 used as the reconstruction loss.

$$\ell_r(x, \hat{x}) = \|x - \hat{x}\|_1 . \tag{4}$$

Reconstruction via multitask learning I

- In the training phase for the multitask network, we minimize the total loss via backpropagation,

$$\min_{\phi(\cdot), f_c(\cdot), f_r(\cdot)} \lambda_c \ell_c\left(y, \hat{y}\right) + \lambda_r \ell_r\left(x, \hat{x}\right), \tag{5}$$

  λ_c and λ_r are the weights to control the loss influence on the multitask network for ℓ_c and ℓ_r, respectively.
- In the decoder, the key element is the deconvolutional layer can be considered as the inverse of a convolutional layer, denoted as $\phi^{\dagger}_{conv}\left(\cdot\right)$, resulting:

$$x = \phi^{\dagger}_{conv}\left(\phi_{conv}\left(x\right)\right). \tag{6}$$

DATASETS AND EXPERIMENTAL SETUP I

Definition: open overall accuracy (OA)

the proportion of corretly classified instances among all instances

$$OA = \frac{\text{Anzahl korrekt klassifizierter Instanzen}}{\text{Gesamtanzahl der Instanzen}} \quad (7)$$

- the value of the open overall accuracy increased from 82.82 to 87.75 for the MDL4OW model compared to the close-world setting indicates that the MDL4OW model performs significantly better than the close-world .

evaluation metric I

Definition: F1-score

a balanced metric combining the Recal and precision .

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

Definition: Recall

indicates the model ability to detect all relevant instances .

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False negative (FP)}} \tag{9}$$

evaluation metric I

Definition: precision

indicates how many of the predicted instances are actualy correct.

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \tag{10}$$

- the value of F1-score decreased from 90.6 to 89.48 for the MDL4OW model compared to the close-world setting indicates a small drop in the balance between precision and recall in the open-world setting.

Definition: Mapping error

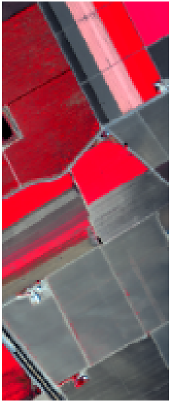
indicates the overall error in prediction the areas between the model's predicted area and the real actual area .

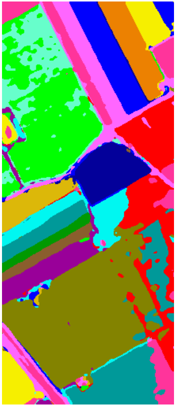
$$\text{Errormax} = 2 \times \left(1 + \frac{A_{gt,C+1}}{\sum_{i=1}^{C} A_{gt,i}}\right) \tag{11}$$

evaluation metric II

- $A_{gt,C+1}$ is the area of unknown class .
- $\sum_{i=1}^{C} A_{gt,i}$ is the total sum of the area of the known classes.
- the value of Mapping error increased from 15.86 to 18.82 for the MDL4OW model compared to the close-world setting,which indicates that the model may have classified known class as unknown due the threshold (=0.5) used .

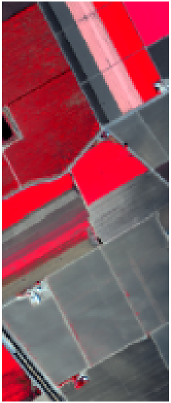
Application: MDL4OW (Multitask Deep Learning) I

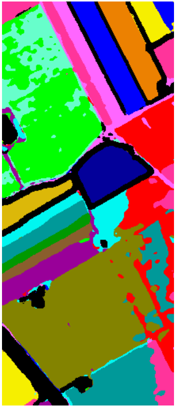




- some of the materials are not represented in the training sample, for exemple, for the right image, the road and the houses between farmlands cannot be classified into any of the known classes.however,the closed model has it classified as combinaition of seen classes, which dosn't reflect the reality .

Application: MDL4OW (Multitask Deep Learning) I





- by using the multitask deep learning the model has the ability to identify the unknown: the road and the hauses between farmlands (masked with black color) are succesfully identified .

Reconstruction via multitask learning I

Reconstruction via multitask learning I

Reconstruction via multitask learning I

ďż˝bung 3.1: ...

...

a) ...

b) ...

c) ...

Lecture 2

Prof. Dr.-Ing. Achim Ibenthal

Hochschule für angewandte Wissenschaft und Kunst
Fakultät für Naturwissenschaften und Technik, Göttingen

SoSe 2024

Ziel: Verständnis Lecture 2

Nutzungsrichtlinie

Hinweise zu Fehlern, fehlenden oder falschen Referenzen, sowie weitere Kommentare und Anregungen sind stets willkommen und werden zur Verbesserung des Vorlesungsmaterials fortlaufend eingearbeitet.

Übersicht

Frame 1 of Subsection 1 of Section 2

...

Übungen

ďż˝bung 4.1: ...

...

a) ...

b) ...

c) ...

Lecture 3

Prof. Dr.-Ing. Achim Ibenthal

Hochschule für angewandte Wissenschaft und Kunst
Fakultät für Naturwissenschaften und Technik, Göttingen

SoSe 2024

Ziel: Verständnis Lecture 3

Nutzungsrichtlinie

Die Inhalte dieser Vorlesung entstammen hauptsďż˝chlich den Quellen [1], [2], [3], die als vorlesungsbegleitendes Material empfohlen werden. Weitere Quellen sind im Text kenntlich gemacht. Darďż˝ber hinaus gehende Inhalte und Bilder sind © A. Ibenthal.

Die Nutzung dieses Dokumentes dient ausschlieďż˝lich der Lehre und ist den Studierenden der HAWK, Fachhochschule Hildesheim, Holzminden, Gďż˝ttingen fďż˝r die Dauer ihrer Immatrikulation gestattet. Die Weitergabe, Verďż˝ffentlichung oder kommerzielle Nutzung – auch nur auszugsweise – ist nicht gestattet.

Hinweise zu Fehlern, fehlenden oder falschen Referenzen, sowie weitere Kommentare und Anregungen sind stets willkommen und werden zur Verbesserung des Vorlesungsmaterials fortlaufend eingearbeitet.

Übersicht

Frame 1 of Subsection 1 of Section 3

...

Übungen

ďż˝bung 5.1: ...

...

a) ...

b) ...

c) ...

Using `settings_common.tex`, `settings_slides.tex`, `settings_slides.tex` and `mybeamer.sty`

Prof. Dr.-Ing. Achim Ibenthal

Hochschule für angewandte Wissenschaft und Kunst
Fakultät für Naturwissenschaften und Technik, Göttingen

SoSe 2024

Goal: Facilitate the application of the `beamer` class, harmonize the colors of block environments and amend box environments providing boxes with harmonized colors but without titles.

Übersicht

General Hints I

- The following files have been created to enhance the flexibility of the beamer class in terms of layout and to provide commands which facilitate its application especially for lectures:
  - settings_common.tex: Things which are common to both slide and script. Loads packages, sets layout and theme colours, adds commands and environments, switches PDF realtime clock for slides on/off, sets author and facility information.
  - settings_slides.tex: slide-specific layout: defines 35 additional color themes, new logo layout, colour control for theme elements.
  - settings_slides.tex: script-specific layout such as page layout and boxes.
  - mybeamer.sty: commands shortcuts for standard text colors, beamerclass control, figures, boxes and math.

General Hints II

- To adapt the overall layout use Section 2 of \format\settings_common.tex.
- To adapt lecture title, personal data and legal notice use Section 3 of \format\settings_common.tex.
- Commands for blocks and boxes are explained in this Section.
- \te{} provides text emphasis, \ce{} character emphasis (e.g. for glossary) and \me{} the marking of menue commands. It is recommended to use those because the layout is optimized for slides and script individually. Defaults are defined in Section 10 of \format\settings_common.tex. A \key{} command is available to add sidenotes to the script.
- beamer increments counters in \section*, \subsection*, \subsubsection*. Therefore use new commands \nosection, \nosubsection, \nosubsubsection instead (not starred).

General Hints III

- `mybeamer.sty` defines many additional commands, which are more or less selfexplanatory when reading `mybeamer.sty`. Examples:
  - Color definitions.
  - commands which operate only in one of the two modes slide and script such as `\slideonly{any text}` and `\scriptonly{any text}`.
  - e.g. `\slup{#}` and `\scrdown{#}` are used for v- and hspaces. Arguments # are passed without unit and interpreted as 'ex'.
  - `\switch{#1}{#2}` separates arguments used in slide mode (#1) and script mode (#2). This comes in quite handy e.g. for figure resizing.
  - Often a `\vspace` is overruled by the TEX auto layout and more or less ignored. Here the commands `\slidebar{#}` and `\scriptbar{#}` are providing the brute force method to enforce spacing.

General Hints IV

- `\fig` facilitates the insertion of figures:

```
\fig{captiontext}
    {fig:<label>}
    {\switch{0.9}{0.7}} % sizes for slide/script
    {figurename}        % some pdf, jpg or png
    \scriptbar{2}       % enforce space in script
```

- There are many more, check it out!

`normalblock` and `normalbox` Environment

normalblock Title

```
\begin{normalblock}{normalblock Title}
    ...
\end{normalblock}
```

- normalblock item

1. normalblock enumerate

```
\begin{normalbox}
    ...
\end{normalbox}
```

- normalbox item

1. normalbox enumerate

exampleblock and examplebox Environment

exampleblock Title

```
\begin{exampleblock}{exampleblock Title}
   ...
\end{normalblock}
```

- examplelblock item

1. exampleblock enumerate

```
\begin{examplebox}
   ...
\end{examplebox}
```

- examplebox item

1. examplebox enumerate

alertblock and alertbox Environment

alertblock Title

```
\begin{alertblock}{alertblock Title}
    ...
\end{alertblock}
```

- alertblock item

1. alertblock enumerate

```
\begin{alertbox}
    ...
\end{alertbox}
```

- alertbox item

1. alertbox enumerate

block and bodybox Environment

The standard `block` environment defined by `beamer` uses theme colors for title, title text, canvas, body text and bullets:

Theme

```
\begin{block}{block Title}
    ...
\end{block}
```

- block item

1. block enumerate

We have added a `bodybox` environment. It's effect is similar to a `beamercolorbox` with `block body` color (but vertically more 'tight'):

```
\begin{bodybox}
    ...
\end{bodybox}
```

- bodybox item

1. body enumerate

Shortcuts

It turns out that in practice typing and error rate can be reduced by using shortcuts for the block and box environments:

```
\nblock{Title}{Body}
\nbox{Body}
\eblock{Title}{Body}
\ebox{Body}
\ablock{Title}{Body}
\abox{Body}
\tblock{Title}{Body}
\tbox{Body}
```

Examples:

```
\nblock
   {Problem}
   {Does dark matter smell?}
```

Problem

Does dark matter smell?

```
\abox{\centering%
      This is important!}
```

This is important!

beamercolorbox Environment

The beamercolorbox environment provides boxes which' canvas is derived from some beamer color. Text and bullet colors are derived from the beamer theme colors (!), however. Better use

- \nbox for blue boxes (n = normal)

- \ebox for green boxes, also used for examples

- \abox for alert boxes

- \tbox for theme color box,

giving you matching bullet colors in presentetion (= beamer) mode and black bullets in article mode anyway.

Definitions \ndef I

The beamer `definition` and `Definition` environments are using template colors for boxes, too. In `settings_common.tex` a new environment for definitions is created which uses blue boxes. This is recommended, since otherwise the audience has to adapt to different box colors in case you are just changing the Theme color.

```
\ndef{def:<label>}{Analogkäse}{
    Das willst Du nicht wirklich wissen!
\filledend}
```

Definition: Analogkäse.

Das willst Du nicht wirklich wissen!

Definitions \ndef II

- We don't provide numbers in presentation (= beamer) mode because
  - they don't mean much to the audience
  - and because they are causing problems when using the beamer overlay specifications such as `\uncover<>`.
- Numbering in script is local to Sections.
- Note that we had to come up with something else than `\definition` or `\Definition` since those were already taken. Thus: `\ndef`
- The 'n' indicates that normal block color is being used.

Satz \nsatz and Theorem \ntheorem

\nsatz (Satz) and \ntheorem (Theorem) essentially are working in the same way as \ndef, except they are producing blue boxes in the script if script boxes are enabled in Section 2 of settings_common.tex.

Theorem: Berti's Sicht.

Die Realität sieht anders aus als die Wirklichkeit.

Satz: Fourier-Transform.

$$S(\omega) = \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt, \tag{12}$$

$$\circ\!\!-\!\!\bullet$$

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega t} d\omega. \tag{13}$$

Korollar \ncor and Lemma \ntheorem

\ncor (Korollar)/\ncor (Corollary) and \nlem (Lemma) do not produce boxes in script mode.

Korollar: Osterhasenpädagogik.

Der Lehrer versteckt das Wissen - die Schüler sollen's suchen.

Corollary: Maturity.

Growing old is mandatory, growing up is optional.

Lemma: Jordan's Lemma.

In der Elektrotechnik gibt es Probleme, bei denen komplexe Zahlen so nützlich sind, wie negative Zahlen für Probleme mit Geld.

The white boxes "$\square$" in script mode are generated using \openend.

Exercises: \Uebung – \Antwort environment pair

For the application of the `\Uebung` – `\Antwort` environment pair for exercises and answers see the LaTeX source code of this template.

You can change the environment title and language of headings in the environment definitions in Section 7 of `settings_common.tex`.

Numbering in script is local to Sections.

Sourcecode Listings I

Source code listings are inserted by using the `listings` package.

This template is configured for `Matlab` M-file language. Use Section 11 of `settings_common.tex` if you wish to change the default.

Sourcecode Listings II

Example:

```
\begin{lstlisting}[title=\cblueb{averagingdemo.m},
                   label=list:average,
                   firstnumber=1,
                   backgroundcolor=\color{ultralightgray}]
% Demo Planarer Tiefpass

%% Bild lesen
b  = imread('zoneplate.tif');

%% Filterung
b5  = imfilter(b, fspecial('average', 5),  'replicate');
b9  = imfilter(b, fspecial('average', 9),  'replicate');
b13 = imfilter(b, fspecial('average', 13), 'replicate');

%% Display
subplot(2,2,1); imshow(b,   []);  title('Original');
subplot(2,2,2); imshow(b5,  []);  title('5 x 5 Filter');
subplot(2,2,3); imshow(b9,  []);  title('9 x 9 Filter');
subplot(2,2,4); imshow(b13, []); title('13 x 13 Filter');
\end{lstlisting}
```

Sourcecode Listings III

```
% Demo Planarer Tiefpass

%% Bild lesen
b   = imread('zoneplate.tif');

%% Filterung
b5  = imfilter(b, fspecial('average', 5),  'replicate');
b9  = imfilter(b, fspecial('average', 9),  'replicate');
b13 = imfilter(b, fspecial('average', 13), 'replicate');

%% Display
subplot(2,2,1); imshow(b,   []);  title('Original');
subplot(2,2,2); imshow(b5,  []);  title('5 x 5 Filter');
subplot(2,2,3); imshow(b9,  []);  title('9 x 9 Filter');
subplot(2,2,4); imshow(b13, []); title('13 x 13 Filter');
```

Literatur I

[1] W.-P. Buchwald,
"Audio-Video-Systeme 2",
Vorlesungsskript FH Braunschweig/Wolfenbüttel, 2005.

[2] B. Wendland,
Fernsehtechnik - Band 1: Grundlagen,
Dr. Alfred Hüthig Verlag GmbH Heidelberg, 1988.

[3] B. Wendland,
Fernsehtechnik - Band 2: Systeme und Komponenten zur Farbbildübertragung,
Dr. Alfred Hüthig Verlag GmbH Heidelberg, 1988.