Hand-designed descriptors vs. pre-trained convolutional networks: a comparison of two strategies for colour texture classification

Francesco Bianconi[1]
School of Mathematics, Computer Science and Engineering
Biomedical Engineering Research Centre
City, University of London
20 November 2018

[1]Department of Engineering, Università degli Studi di Perugia, Perugia, Italy

Outline

Outline

About the speaker

About the speaker (1/2)

- **Contacts**
  - Department of Engineering, Università degli Studi di Perugia Via G. Duranti 93, 06125 Perugia (Italy)
  - e-mail: bianco@ieee.org
- **Position**: *Professore Associato* (eqv. UK = Senior Lecturer)
  - Currently *Academic Visitor* in the School of Mathematics, Computer Science & Engineering at City, University of London
- **Background** in Industrial Engineering (MEng Mechanical Engineering, PhD Computer-aided Design), but gradually moved to Computer Science

About the speaker (2/2)

- **Research interests** in the area of image processing, with particular focus on:
  - Colour and texture analysis
    - Theory (*mathematics of textures*)
    - Industrial applications (*surface sorting* and *grading*, *defect detection*)
    - Biomedical applications (*histology*, *radiomics*)

The University of Perugia

- Established in 1308;
- About 26,000 students;
- Sixteen departments:
  - Agricultural Sciences; Ancient and Modern Literatures, Languages and Cultures; Chemistry, Biology and Biotechnologies; Civil and Environmental Engineering; Economics; Engineering; Experimental Medicine; Law; Mathematics and Computer Science; Medicine; Philosophy, Human and Social Sciences; Pharmacy; Physics and Geology; Political Sciences; Surgery and Biomedicine; Veterinary.
- Ranked 1st for the fourth year in a row among the Italian universities with a student population between 20,000 and 40,000 (CENSIS, 2018)

The city of Perugia


- A well-preserved medieval town
- Population of about 160,000
- Quiet and safe but culturally active
- Good environment for students

Umbria: ‘The green heart of Italy’

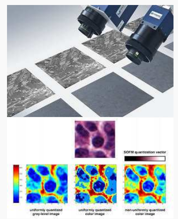
Colour and texture

Colour and texture

- Colour and texture are two fundamental **visual stimuli** that determine most of the appearance of materials, objects and scenes
  - **Colour**: property of **a point** or **an area**
  - **Texture**: property of **an area**, related to the spatial variation of the visual appearance
- The ability to describe colour and texture plays a fundamental role in many computer-vision tasks, e.g.:
  - **Classification**
  - **Segmentation**
  - **Content-Based Image Retrieval**

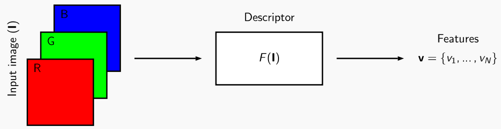
Applications of colour texture analysis

- Colour texture analysis is of vital importance in a wide range of applications, as for instance:
  - Industrial
    - Surface grading
    - Surface inspection
    - Defect detection
  - Biomedical
    - Computer-assisted diagnosis
    - Tissue classification



Objective of colour texture analysis

- To extract **meaningful** and **discriminative** parameters from the input images (**features**)
  - The images are represented through their features
  - The features usually are the input to some kind of classifier to perform tasks as classification, segmentation or retrieval



- $N \rightarrow$ dimensionality of the descriptor

Strategies to colour texture analysis

- **Traditional approach**: '*hand-designed*' (also referred to as '*hand-crafted*' or '*engineered*') features
  - Dominant from the **early 70's** (Haralick et al., 1973) up until not long ago
  - Defined by hand (**a priori**, **unlearnt**)
- **Convolutional Neural Networks** ('*Deep Learning*')
  - Increasing attention in the last few years
  - Parameters estimated by training (**a posteriori**, **learnt**)
  - Unprecedented performance in object, scene and face recognition
  - Not quite clear yet if/how this model can scale well to textures

Hand-designed descriptors vs. Deep Learning: an overview

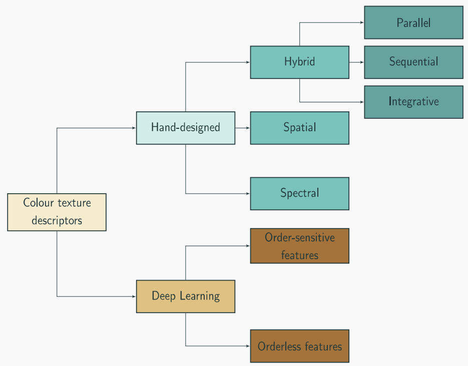
Hand-designed

- Mostly defined **a priori** (little or no training needed)
- Relatively **intuitive** / **interpretable** features
- Low computational demand
- **Unconstrained** input field
- Accuracy can be very **dataset-dependent**
- Need to **find the right method** for any specific task

Deep learning

- Based on **modules** and sets of **free parameters** (to be determined by training)
- Features **difficult to interpret**
- Require **large datasets** for training and dedicated hardware
- **Constrained** (fixed) input field
- **Unprecedented accuracy** in many tasks
- **Generalizable** results

A taxonomy of colour texture descriptors



Spectral methods (colour descriptors)

- Statistical distribution of **colour only – no spatial information**
  - Rather resilient to rotation, changes in scale and/or viewpoint
  - Sensitive to illumination changes

Method	Acronym	N	Ref./Year
Mean of each channel	Mean	3	Kukkonen et al. (2001)
Mean and std. dev. of each channel	Mean+Std	6	Kukkonen et al. (2001)
Mean and $2^{th} - 5^{th}$ moms. of each channel	Mean+Moms	15	López et al. (2008)
Quartiles of each channel	Quartiles	9	Niskanen et al. (2001)
256-bin marginal histogram of each channel	Marginal-Hists-256	768	Pietikainen et al. (1996)
10-bin joint colour histogram	Full-Hist-10	1000	Swain and Ballard (1991)

Spatial methods (texture descriptors) (1/2)

- Computed on the **luminance** channel – **colour is discarded**
  - Rather resilient to illumination changes
  - Sensitive to rotation, changes in scale and/or viewpoint (some degree of invariancecan be achieved)

Method	Acronym	N	Ref./Year
Completed Local Binary Patterns	CLBP	324	Guo et al. (2010)
Gradient-Based Local Binary Patterns	GLBP	108	Jin et al. (2004)
Improved Local Binary Patterns	ILBP	213	Jin et al. (2004)
Local Binary Patterns	LBP	108	Ojala et al. (2002)
Local Ternary Patterns	LTP	216	Tan and Triggs (2010)
Texture Spectrum	TS	2502	He and Wang (1990)
Grey-level Co-occurrence Matrices	GLCM	60	Haralick et al. (1973)
Gabor features	Gabor	70	Manjunath and Ma (1996)
Image Patch-Based Classifier	IPBC-J	4096	Varma and Zisserman (2009)
Histograms of Oriented Gradients	HOG	768	Dalal (2005)
Dense SIFT	SIFT	4096	Lowe (2004)
VZ Classifier	VZ-MR8	4096	Varma and Zisserman (2005)
Wavelet statistical and co-occ. feats.	WSF+WCF	84	Arivazhagan and Ganesan (2003)

Spatial methods (texture descriptors) (2/2)

- *GLCM*, *HOG*, *IPBC-J*, *LBP-variants* and *SIFT*: concatenation of features computed at resolution 1px, 2px and 3px
- *Gabor*: bank of five freqs. and seven ornts.; with and without DFT normalisation for rotation-invariance; with and without contrast normalisation
- *LBP-variants*: rotation-invariant features ('ri' version)
- *Wavelets*: statistical and co-occurrence features of a three-level decomposition ('haar' and 'bi-orhtogonal')
- *IPBC-J*, *SIFT* and *VZ-MR8*: features aggregated over a pre-defined (external) dictionary of 4096 visual words

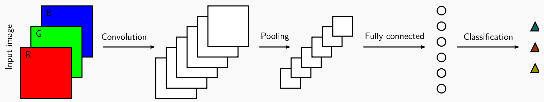
Hybrid methods (colour texture descriptors)

- Combine colour and texture in different ways (Palm, 2004)
  - **Parallel**: extract colour and texture separately then concatenate
  - **Sequential**: convert to single-channel then extract texture features
  - **Integrative**: extract texture features from each colour channel and/or couples of channels

Hybrid methods	Acronym	N
Improved Opponent Colour Local Binary Patterns	IOCLBP	1287
Opponent Colour Local Binary Patterns	OCLBP	648
Local Binary Patterns + Local Colour Contrast	LBP+LCC	876
Local Colour Vector Binary Patterns	LCVBP	432
Integrative Co-occurrence Matrices	ICM	360
Opponent Gabor features	OppGabor	630

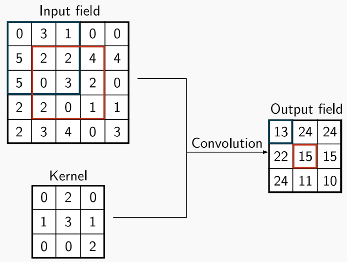
Convolutional Neural Networks: overall architecture

- Based on sets of modules that can be combined together, e.g.: **convolutional layers**, **pooling layers**, **rectifying units**, and **fully-connected layers**
- Some modules (**learnable layers**) include parameters that need to be set by training



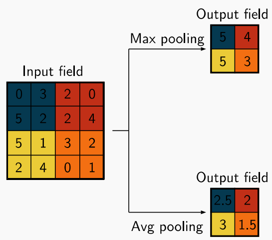
Convolutional layers

- Banks of linear filters
  - Structure of the kernel defined a priori (e.g. 3px × 3px)
  - Weights learned by training



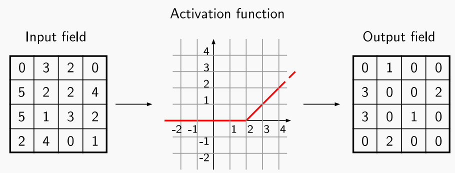
Pooling layers

- Reduce the spatial size of the representation



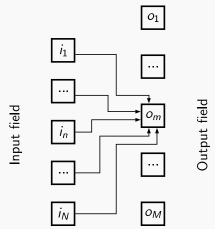
Rectifying units

- Implement some kind of activation function (e.g.: piecewise linear, hyperbolic tangent or sigmoidal)



Fully-connected layers

- Each value of the output field depends on all the values of the input field
  - Can be implemented as convolutional layer where the kernel is the same size as the input field



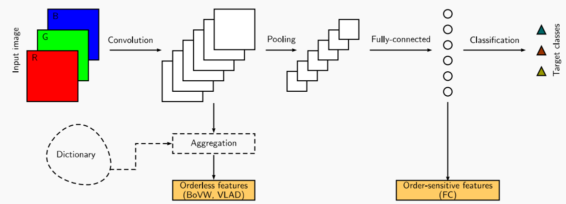
Approaches to convolutional networks

- **Full training**: the network is trained from scratch, all the parameters need to be 'learned'
  - Requires large datasets of images
  - Computationally expensive
- **Transfer learning**: take a network trained for one task (e.g. face recognition) and use it *as is* for another task (e.g. texture classification)
  - No training, little overhead
  - Computationally cheap
- **Fine tuning**: take a network trained for one task and optimise some of its parameters for another task
  - Requires medium-sized datasets
  - Computationally not so expensive

Pre-trained CNN as feature extractors (1/2)

- Common approaches (Cimpoi et al., 2016):
  - **Order-sensitive**: take the output of the **last fully-connected layer** ('**FC**', *global* representation)
  - **Orderless**: Take the output of the **last convolutional layer** and **aggregate** it over a pre-defined dictionary of visual words (*local* representation)
    - Bag of visual words ('**BoVW**')
    - Vector of locally-aggregated descriptor encoder ('**VLAD**')

Pre-trained CNN as feature extractors (2/2)



Convolutional Neural Networks: pre-trained models (1/2)

- Eleven pre-trained models (10 trained for object recognition, one for face recognition)
- Input field of 224px × 224px

Pre-trained model	No. of features	Ref./Year
GoogLeNet-FC	1024	Szegedy et al. (2015)
GoogLeNet-BoVW	1024	Szegedy et al. (2015)
Caffe-Alex-FC	4096	Krizhevsky et al. (2012)
Caffe-Alex-BoVW	4096	Krizhevsky et al. (2012)
Caffe-Alex-VLAD	4224	Krizhevsky et al. (2012)
ResNet-50-FC	2048	He et al. (2016)
ResNet-50-BoVW	2048	He et al. (2016)
ResNet-101-FC	2048	He et al. (2016)
ResNet-101-BoVW	2048	He et al. (2016)
ResNet-152-FC	2048	He et al. (2016)
ResNet-152-BoVW	2048	He et al. (2016)
VGG-F-FC	4096	Chatfield et al. (2014)
VGG-F-BoVW	4096	Chatfield et al. (2014)
VGG-F-VLAD	4096	Chatfield et al. (2014)

Convolutional Neural Networks: pre-trained models (2/2)

Pre-trained model	No. of features	Ref./Year
VGG-M-FC	4096	Chatfield et al. (2014)
VGG-M-BoVW	4096	Chatfield et al. (2014)
VGG-M-VLAD	4096	Chatfield et al. (2014)
VGG-S-FC	4096	Chatfield et al. (2014)
VGG-S-BoVW	4096	Chatfield et al. (2014)
VGG-S-VLAD	4096	Chatfield et al. (2014)
VGG-VD-16-FC	4096	Simonyan and Zisserman (2015)
VGG-VD-16-BoVW	4096	Simonyan and Zisserman (2015)
VGG-VD-16-VLAD	4096	Simonyan and Zisserman (2015)
VGG-VD-19-FC	4096	Simonyan and Zisserman (2015)
VGG-VD-19-BoVW	4096	Simonyan and Zisserman (2015)
VGG-VD-19-VLAD	4096	Simonyan and Zisserman (2015)
VGG-Face-FC	4096	Parkhi et al. (2015)
VGG-Face-BoVW	4096	Parkhi et al. (2015)
VGG-Face-VLAD	4096	Parkhi et al. (2015)

Experiments

Objectives

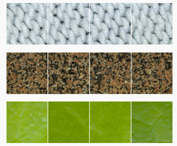
- To investigate the effectiveness of **hand-designed descriptors** vs. **pre-trained convolutional networks** for colour texture classification
    - To evaluate accuracy with **stationary** and **non-stationary** textures
        - Under **ideal** conditions (steady acquisition settings)
        - Under **realistic** conditions (with changes in illumination, rotation, scale and/or viewpoint)

Colour texture datasets

- **Fifty-eight datasets** of colour textures from **31** image repositories
  - Generic textures (materials and scenes)
  - Biomedical textures
- **Ten experimental conditions** depending on:
  - The stationariness of the texture – i.e: stationary vs. non-stationary
  - The presence/absence of variations in the imaging conditions such as changes in **illumination**, **rotation**, **viewpoint** and/or **scale**

Stationary vs. non-stationary textures

- **Stationary texture**: '*an image which contains a single type of texture . . . and so its local statistical properties are the same everywhere*' (Petrou and García Sevilla, 2006)

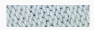
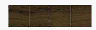
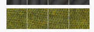
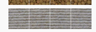
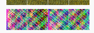
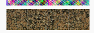
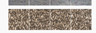
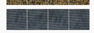
Stationary

Non-stationary


Variations in the imaging conditions

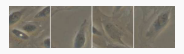
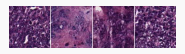
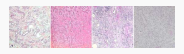
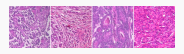
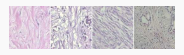
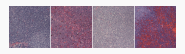
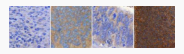
Symbol	Imaging conditions	Abb.
None	Steady	N
	Variable illumination	I
	Rotation	R
	Changes in scale	S
	Change in viewpoint	–
–	Multiple variations	M

Group #1: Stationary textures acquired under steady imaging conditions (1/2)

- Twenty-five datasets

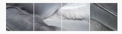
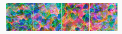
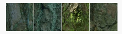
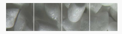
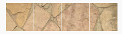
Name	Var.	Sample images	Name	Var.	Sample images
ALOT (95)	None		Parquet (38)	None	
CBT (100)	None		PlantLeaves (20)	None	
Drexel (18)	None		RawFoot (68)	None	
KylbergSintorn (25)	None		STex (202)	None	
MBT (120)	None		USPTex (137)	None	
MondialMarmi (25)	None		VisTex (89)	None	
Outex (192)	None		VxC_TSG (42)	None	

Group #1: Stationary textures acquired under steady imaging conditions (2/2)

Name	Var.	Sample images	Name	Var.	Sample images
BioMediTechRPE (4)	None		Kather (8)	None	
BreakHis40X (2)	None		LiverAgeing (4)	None	
BreakHis100X (2)	None		LiverGender-AL (2)	None	
BreakHis200X (2)	None		LiverGender-CR (2)	None	
BreakHis400X (2)	None		Lymphoma (3)	None	
Epistroma (2)	None				

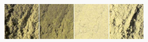
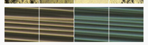
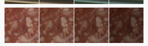
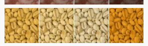
Group #2: Non-stationary textures acquired under steady imaging conditions

- Eight datasets

Name	Variations	Sample images
ALOT (40)	None	
ForestSpecies (112)	None	
MBT (34)	None	
NewBarkTex (6)	None	
Outex (59)	None	
STex (141)	None	
USPTex (33)	None	
VisTex (78)	None	

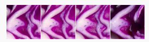
Group #3: Stationary textures acquired under variable illumination

- Five datasets

Name	Variations	Sample images
ALOT-95-S-I	5×	
Outex-192-S-I	3×	
RawFooT-68-S-I1	4×	
RawFooT-68-S-I2	6×	
RawFooT-68-S-I3	3×	

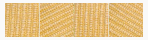
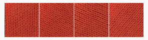
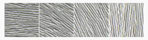
Group #4: Non-stationary textures acquired under variable illumination

- Two datasets

Name	Variations	Sample images
ALOT-40-NS-I	5×	
Outex-59-NS-I	3×	

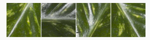
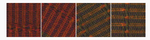
Group #5: Stationary textures with rotation

- Four datasets

Name	Variations	Sample images
ALOT-95-S-R	4×	
KylbergSintorn-25-S-R	4×	
MondialMarmi-25-S-R	4×	
Outex-192-S-R	4×	

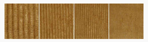
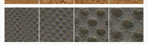
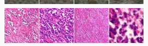
Group #6: Non-stationary textures with rotation

- Two datasets

Name	Variations	Sample images
ALOT-40-NS-R	4×	
Outex-59-NS-R	4×	

Group #7: Stationary textures with variations in scale

- Four datasets

Name	Variations	Sample images
KTH-TIPS-10-S-S	9×	
KTH-TIPS2b-11-S-S	9×	
Outex-192-S-S	4×	
BreakHis-2-S-N	4×	

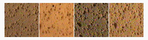
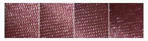
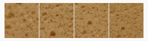
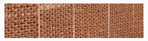
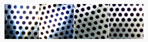
Group #8: Non-stationary textures with variations in scale

- One dataset

Name	Variations	Sample images
Outex-59-NS-S	4×	

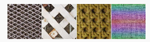
Group #9: Stationary textures acquired under multiple variations

- Six datasets

Name	Variations	Sample images
CUReT-61-S-M		
Fabrics-1968-S-M		
KTH-TIPS-10-S-M		
KTH-TIPS2b-11-S-M		
LMT-94-S-M		
RDAD-27-S-M		

Group #10: Non-stationary textures acquired under multiple variations

- One dataset

Name	Variations	Sample images
DTD-47-NS-M		

Experimental set-up

- Supervised image classification (10 experiments, one for each experimental condition)
- Accuracy estimation based on split sample validation with stratified sampling:
  - A fraction f of the samples of each class for training the classifier (**train set**)
  - The remaining fraction $(1 - f)$ for accuracy estimation (**test set**)
  - Results averaged of 100 random splits into train and test set
  - $f = 1/8$
- Classifier: nearest-neighbour rule with L_1 ('cityblock') distance

Evaluation protocol

- For each dataset a **ranking** was computed through an **all-play-all** tournament
  - *Win*: +1 point
  - *Draw*: 0 points
  - *Loss*: -1 point
- Win (loss) assigned if the accuracy of one descriptor was significantly higher (lower) than the other's
  - Non-parametric Wilcoxon-Mann-Whitney rank sum test ($\alpha = 0.05$)
- Across-dataset ranking obtained by summing up the points that every descriptor obtained in each dataset

Results and discussion

Overall ranking: ten best CNN-based descriptors

Descriptor	Average accuracy*	Average rank*	Overall position
ResNet-50-FC	83.6	63.1	**1**
ResNet-101-FC	83.2	62.5	**2**
ResNet-152-FC	83.1	61.8	**3**
VGG-VD-16-FC	80.3	58.0	**4**
VGG-VD-19-FC	79.9	56.7	**5**
VGG-M-FC	78.6	56.5	**6**
VGG-S-FC	78.4	55.8	**7**
VGG-F-FC	77.6	54.5	**8**
VGG-S-VLAD	75.8	51.4	**10**
Caffe-Alex-FC	75.8	50.8	**11**

* Results averaged over the 10 datasets

Overall ranking: ten best hand-designed descriptors

Descriptor	Average accuracy	Average rank	Overall position
IOCLBP	73.9	51.8	**9**
OCLBP	72.3	47.7	**13**
ICM^{DFT}	69.9	42.9	**16**
LCVBP	69.6	42.9	**17**
Full-Hist-10	69.3	40.7	**19**
ICM	68.4	38.7	**21**
SIFT-BoVW	68.8	38.1	**23**
Marginal-Hists-256	66.4	38.1	**24**
ILBP	67.6	38.0	**25**
LBP+LCC	67.3	37.6	**26**

Experiments #1 and #2: steady conditions

Type of	Hand-designed		CNN-based	
texture	Competitive	Best methods	Competitive	Best methods
Stationary	Yes	IOCLBP (91.9%) OCLBP (91.5%)	Yes	ResNet-50-FC (91.4%) ResNet-101-FC (90.5%)
Non-stationary	Nope	LCVBP (82.1%) IOCLBP (78.6%)	Yes	ResNets-50-FC (87.9%) ResNets-101-FC (87.1%)

*Figures between parentheses indicate average accuracy

**Competitive yes/nope indicates a difference > 5pp between the best hand-designed descriptor and the best CNN-based method

Experiments #3 and #4: variable illumination

Type of texture	Hand-designed Competitive	Hand-designed Best methods	CNN-based Competitive	CNN-based Best methods
Stationary	Nope	SIFT (87.1%) IOCLBP (84.4%)	Yes	ResNet-50-FC (94.8%) ResNet-101-FC (94.4%)
Non-stationary	Nope	SIFT (72.4%) IOCLBP (71.6%)	Yes	ResNet-50-FC (84.8%) ResNet-101-FC (84.0%)

Experiments #5 and #6: rotation

Type of texture	Hand-designed Competitive	Hand-designed Best methods	CNN-based Competitive	CNN-based Best methods
Stationary	Yes	OCLBP (97.2%) IOCLBP (96.8%)	Yes	ResNet-50-FC (96.1%) ResNet-101-FC (95.9%)
Non-stationary	Yes	Full-Hist-10 (84.4%) IOCLBP (82.4%)	Yes	ResNet-50-FC (84.6%) ResNet-50-FC (83.9%)

Experiments #7 and #8: changes in scale

Type of texture	Hand-designed Competitive	Hand-designed Best methods	CNN-based Competitive	CNN-based Best methods
Stationary	Nope	Marginal-Hists-256 (76.2%) Full-Hist-10 (74.9%)	Yes	ResNet-101-FC (86.0%) ResNet-152-FC (85.5%)
Non-stationary	Yes	Full-Hist-10 (73.3%) Marginal-Hists-256 (73.2%)	Yes	ResNet-50-FC (70.4%) ResNet-152-FC (69.5%)

Experiments #9 and #10: multiple variations ('in the wild')

Type of texture	Hand-designed Competitive	Hand-designed Best methods	CNN-based Competitive	CNN-based Best methods
Stationary	Nope	IOCLBP (67.9%) OCLBP (66.3%)	Yes	ResNet-152-FC (83.6%) ResNet-101-FC (83.4%)
Non-stationary	Nope	SIFT (31.6%) VZ-MR8 (27.8%)	Yes	ResNet-50-FC (60.8%) ResNet-152-FC (60.4%)

Computational demand

- Hand-designed descriptors approximately as fast as CNN-based methods in feature extraction
- CNN-based methods slower in classification (longer feature vectors)

Discussion (1/2)

- Results generally favourable to the **CNN-based** methods
  - Consistent with Napoletano (2017)
- Features from **fully-connected layers** better than features from convolutional layers
  - In line with Cusano et al. (2016)
  - Significantly different from Cimpoi et al. (2016)
- Colour variants of LBP (IOCLBP, OCLBP) best among the hand-crafted methods

Discussion (2/2)

Hand-designed methods

- Good with **stationary** textures and **steady** imaging conditions
- **Non-competitive** with **non-stationary** textures and **large variations** in the imaging conditions
- Accuracy **strongly dependent** on the experimental conditions: **no all-purpose method**, need to find the right descriptor for each specific task

CNN-based descriptors

- Good accuracy both with **stationary** and **non-stationary** textures
- **Resilient** to changes in the imaging conditions
- **Stable** performance across different experimental conditions (ResNet always best model)

Conclusions

Conclusions

- The use of pre-trained convolutional neural networks as generic feature extractors is a **viable** approach to colour texture classification
- **Pre-trained CNNs**
  - Good with **non-stationary** textures and in the presence of **multiple variations** in the imaging conditions
- **Hand-designed descriptors**
  - Competitive with **stationary textures** and **steady** imaging conditions

References & further readings (1/6)

- S. Arivazhagan and L. Ganesan. "Texture classification using wavelet transform". In: *Pattern Recognition Letters* 24.9-10 (June 2003), pp. 1513–1521.
- CENSIS. *Le classifiche degli atenei statali (2018/2019)*. Available online at `http://www.censis.it/17?shadow_pubblicazione=120579`. Last accessed on Sep.. 4, 2018. 2018.
- K. Chatfield et al. "Return of the devil in the details: Delving deep into convolutional nets". In: *Proceedings of the British Machine Vision Conference*. Nottingham, United Kingdom, Sept. 2014.
- M. Cimpoi et al. "Deep Filter Banks for Texture Recognition, Description, and Segmentation". In: *International Journal of Computer Vision* 118.1 (2016), pp. 65–94.
- C. Cusano, P. Napoletano, and R. Schettini. "Evaluating color texture descriptors under large variations of controlled lighting conditions". In: *Journal of the Optical Society of America A: Optics and Image Science, and Vision* 33.1 (2016), pp. 17–30.

References & further readings (2/6)

- B. Dalal N. end Triggs. "Histograms of oriented gradients for human detection". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. Vol. 1. Art. no. 1467360. San Diego, USA, June 2005, pp. 886–893.
- Z. Guo, L. Zhang, and D. Zhang. "A completed modeling of local binary pattern operator for texture classification". In: *IEEE Transactions on Image Processing* 19.6 (2010), pp. 1657–1663.
- R. M. Haralick, K. Shanmugam, and I. Dinstein. "Textural Features for Image Classification". In: *IEEE Transactions on Systems, Man, and Cybernetics* 3.6 (1973), pp. 610–621.
- D.-C. He and L. Wang. "Texture Unit, Texture Spectrum, And Texture Analysis". In: *IEEE Transactions on Geoscience and Remote Sensing* 28.4 (1990), pp. 509–512.
- K. He et al. "Deep residual learning for image recognition". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Las Vegas, United States, Jan. 2016, pp. 770–778.

References & further readings (3/6)

- H Jin et al. "Face detection using improved LBP under Bayesian framework". In: *Proceedings of the 3rd International Conference on Image and Graphics*. Hong Kong, China, Dec. 2004, pp. 306–309.
- A. Krizhevsky, I. Sutskever, and G.E. Hinton. "ImageNet classification with deep convolutional neural networks". In: *Proceedings of the 26th Annual Conference on Neural Information Processing Systems*. Lake Tahoe, United States, Dec. 2012.
- S. Kukkonen, H. Kälviäinen, and J. Parkkinen. "Color features for quality control in ceramic tile industry". In: *Optical Engineering* 40.2 (Feb. 2001), pp. 170–177.
- F. López et al. "Performance evaluation of soft color texture descriptors for surface grading using experimental design and logistic regression". In: *Pattern Recognition* 41.5 (May 2008), pp. 1761–1772.
- D. Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". In: *International Journal of Computer Vision* 60.2 (2004), pp. 91–110.

References & further readings (4/6)

- B. S. Manjunath and W. Y. Ma. "Texture Features for Browsing and Retrieval of Image Data". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18.8 (1996), pp. 837–841.
- P. Napoletano. "Hand-Crafted vs Learned Descriptors for Color Texture Classification". In: *Proceedings of the 6th Computational Color Imaging Workshop (CCIW'17)*. Ed. by S. Bianco et al. Vol. 10213. Lecture Notes in Computer Science. Milan, Italy: Springer, Mar. 2017, pp. 259–271.
- M. Niskanen, O. Silvén, and H. Kauppinen. "Color and texture based wood inspection with non-supervised clustering". In: *Proceedings of the 12th Scandivanian Conference on Image Analysis (SCIA 2001)*. Bergen, Norway, June 2001, pp. 336–342.
- T. Ojala, M. Pietikäinen, and T. Mäenpää. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.7 (2002), pp. 971–987.
- C. Palm. "Color texture classification by integrative Co-occurrence matrices". In: *Pattern Recognition* 37.5 (2004), pp. 965–976.

References & further readings (5/6)

- O. Parkhi, A. Vedaldi, and A Zisserman. "Deep face recognition". In: *Proceedings of the British Machine Vision Conference*. Swansea, United Kingdom, Sept. 2015.
- M. Petrou and P. García Sevilla. *Image Processing. Dealing with Texture.* Wiley Interscience, 2006.
- M. Pietikainen et al. "Accurate color discrimination with classification based on feature distributions". In: *Proceedings of the International Conference on Pattern Recognition (ICPR)*. Vol. 3. Art. no. 547285. Vienna, Austria, Aug. 1996, pp. 833–838.
- K. Simonyan and A. Zisserman. "Very deep convolutional networks for large-scale image recognition". In: *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. San Diego, USA, May 2015.
- M.J. Swain and D.H. Ballard. "Color indexing". In: *International Journal of Computer Vision* 7.1 (Nov. 1991), pp. 11–32.

References & further readings (6/6)

- C. Szegedy et al. "Going deeper with convolutions". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Boston, United States, June 2015, pp. 1–9.
- X. Tan and B. Triggs. "Enhanced local texture feature sets for face recognition under difficult lighting conditions". In: *IEEE Transactions on Image Processing* 19.6 (2010), pp. 1635–1650.
- M. Varma and A. Zisserman. "A statistical approach to material classification using image patch exemplars". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31.11 (2009), pp. 2032–2047.
- M. Varma and A. Zisserman. "A statistical approach to texture classification from single images". In: *International Journal of Computer Vision* 62.1-2 (Apr. 2005), pp. 61–81.