

Gesichtserkennung mittels Deep Learning

Jan Kaßel

26. Januar 2018

Problemseminar „Deep Learning“
Wintersemester 2017/2018
Universität Leipzig

- Bilderkennung
- Danach Spracherkennung mittels Recurrent Neural Networks
- Convolutional Neural Networks
 - Kantenerkennung, semantische Features
 - Abstrakte Repräsentationen

Deep Learning Identity- Preserving Face Space

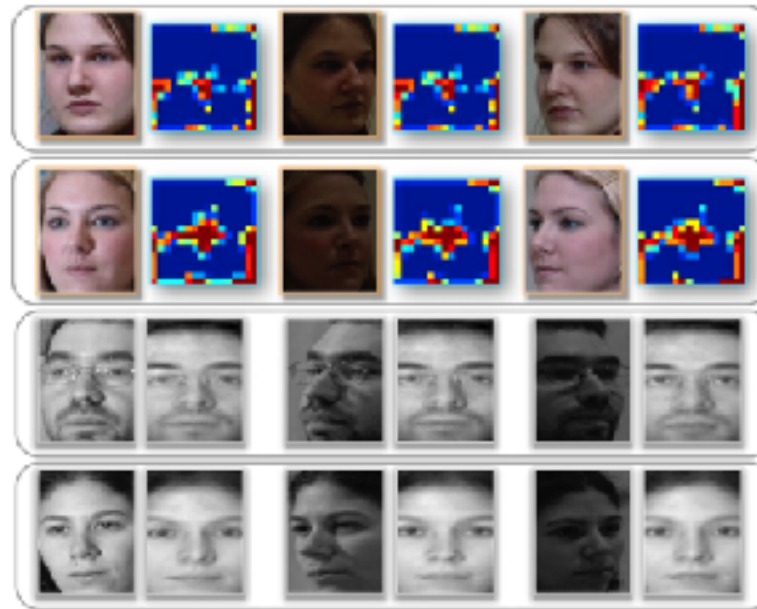
Zhu et al. ICCV, 2013

- Gesichtsrepräsentation mittels
Deep Neural Network
- Stabil über Varianzen hinweg
- Rekonstruktion des Gesichts

3

- 5 Jahre alt
- Repräsentation, nicht Vergleich
 - Notwendig für Vergleich
 - Merkmale im Gesicht enkodieren, (un)vergleichbar machen
 - Stabil über verschiedene Blickwinkel
 - Rekonstruktion in neutraler Ansicht

Deep Learning Identity- Preserving Face Space



4

- Person pro Reihe
 - Oben encodierte Features
 - unten wiederhergestelltes Gesicht, Canonical View

Inhalt

- **Motivation**
- **Ansätze**
- „Identity-Preserving Face Space“
 - Architektur
 - Training
 - Auswertung

5

- Motivation: Anwendung aus der Gegenwart
- Ansätze: Methoden, die zum heutigen Stand geführt haben
- Paper selbst
 - Architektur, dessen Training
 - Performance im Vergleich

Anwendungen



Identifikation



Überwachung

6

- Viele populäre Anwendungen
- iPhone 10
 - Face ID: neues System, mittels Gesichtserkennung Gerät sichern
 - Skaliert
 - Stabil
 - Video Surveillance
 - China
 - Problematiken mit Stabilität



- Face ID
- Funktioniert im Dunkeln, auch mit Brille
- Projiziert Infrarot-Matrix auf das Gesicht
- Scanned es
- Erstellt 3D-Repräsentation



- Face ID
- Funktioniert im Dunkeln, auch mit Brille
- Projiziert Infrarot-Matrix auf das Gesicht
- Scanned es
- Erstellt 3D-Repräsentation



- Wall Street Journal
 - Bericht über Überwachung in China
 - Schwierigkeiten
 - Winkel
 - Belichtungen



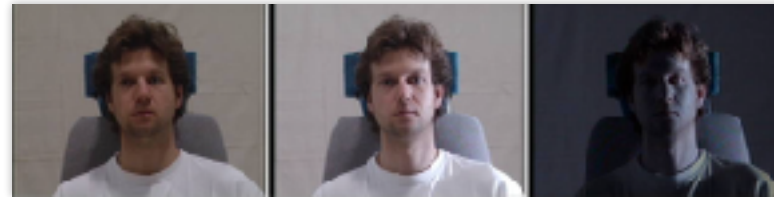
- Wall Street Journal
 - Bericht über Überwachung in China
 - Schwierigkeiten
 - Winkel
 - Belichtungen

Problematiken

- Pose



- Belichtung



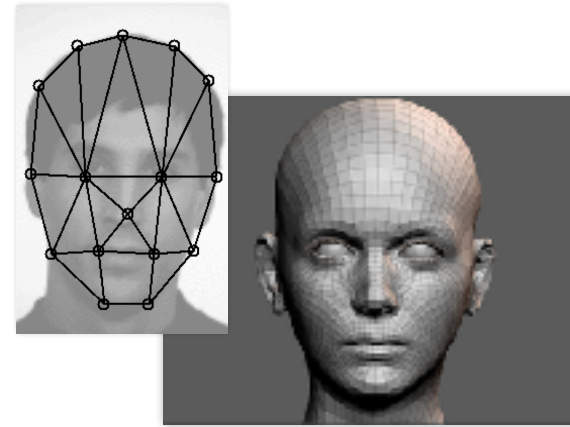
- Fragestellung
 - Gesichter erkennen trotz
 - Pose
 - Belichtung

Inhalt

- Motivation
- **Ansätze**
- „Identity-Preserving Face Space“
 - Architektur
 - Training
 - Auswertung

Ansätze

- 2D- vs. 3D-basierte Methoden
- „traditionelle“ Deskriptoren vs. Deep Learning



- 2D: Einfach zu bekommen
 - 3D: Schwieriger, berechnen? Aber error-prone
- Traditionell: „hand-crafted“, mathematisch
 - Deep Learning: Komplexe Eingaben, komplexe Deskriptoren

Gabor Descriptor

Wiskott et al., 1997



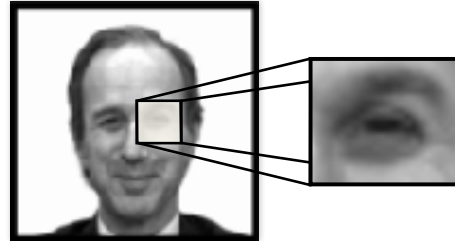
Eingabe

12

- Gabor Filter: 2D-Texturanalyse
 - Frequenz- und Orientierungsanalyse
 - Ähnlich dem menschlichen Sichtsystem
- Kombination in Jets
- Jets werden mit Bunch Graph assoziiert
 - Flexible Zuweisung

Gabor Descriptor

Wiskott et al., 1997



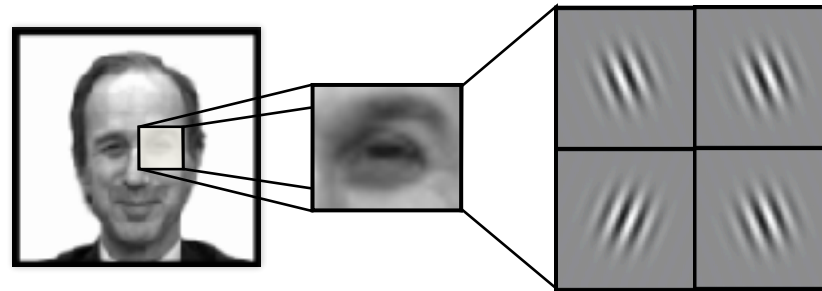
Eingabe

12

- Gabor Filter: 2D-Texturanalyse
 - Frequenz- und Orientierungsanalyse
 - Ähnlich dem menschlichen Sichtsystem
- Kombination in Jets
- Jets werden mit Bunch Graph assoziiert
 - Flexible Zuweisung

Gabor Descriptor

Wiskott et al., 1997



Eingabe

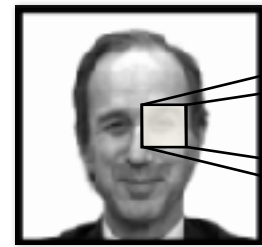
Jet

12

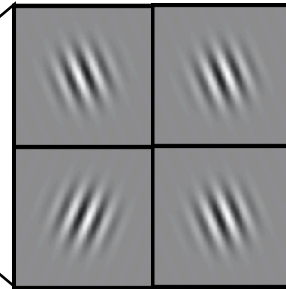
- Gabor Filter: 2D-Texturanalyse
 - Frequenz- und Orientierungsanalyse
 - Ähnlich dem menschlichen Sichtsystem
- Kombination in Jets
- Jets werden mit Bunch Graph assoziiert
 - Flexible Zuweisung

Gabor Descriptor

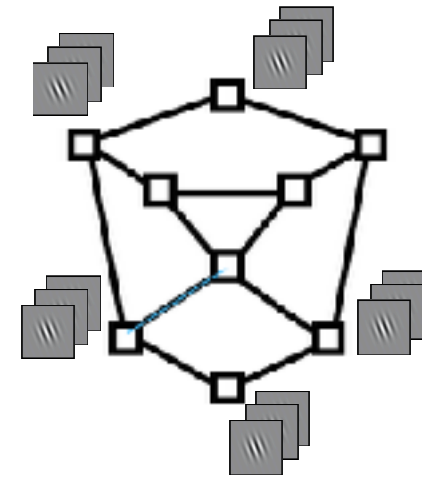
Wiskott et al., 1997



Eingabe



Jet



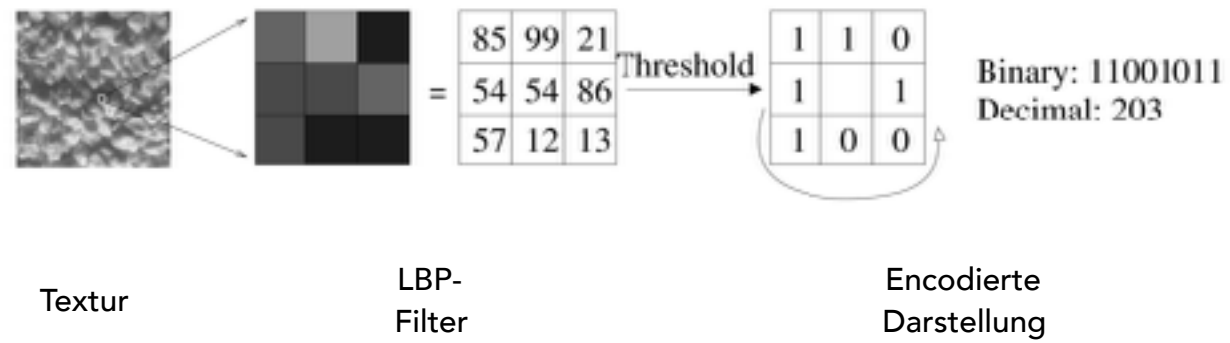
Bunch Graph

12

- Gabor Filter: 2D-Texturanalyse
 - Frequenz- und Orientierungsanalyse
 - Ähnlich dem menschlichen Sichtsystem
- Kombination in Jets
- Jets werden mit Bunch Graph assoziiert
 - Flexible Zuweisung

Local Binary Pattern Descriptor

Ahonen et al., 2006



13

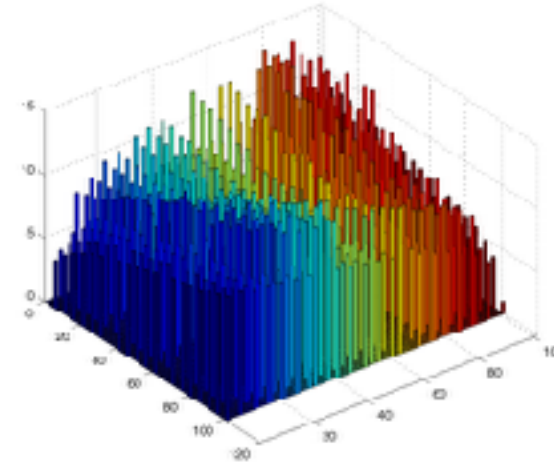
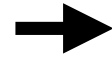
- LBP: Klassischer Textur-Deskriptor
- Filtern mit kleinem Ausschnitt
- Encodieren

Local Binary Pattern Descriptor

Ahonen et al., 2006

1	1	0
1		1
1	0	0

Binary: 11001011
Decimal: 203



Encodierte
Darstellung

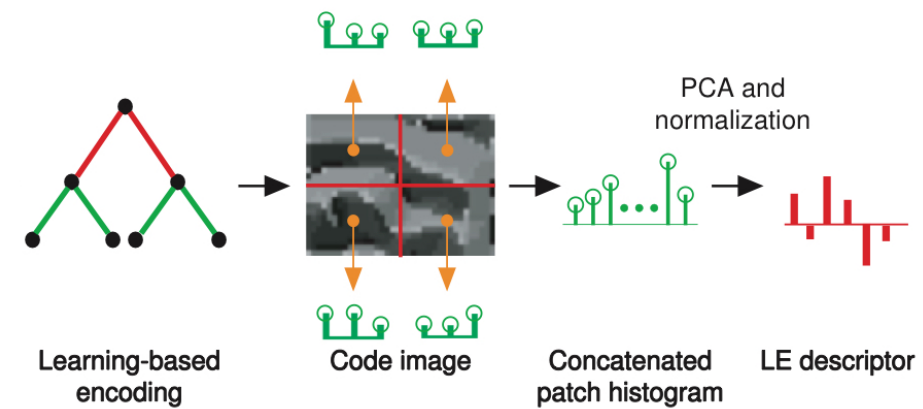
Deskriptor

14

- Dann Zuweisung zu bestimmten Gesichtsregionen, Art Heatmap
 - lokale Deskriptoren zu global Deskriptor Kombiniert

Learning-Based Descriptor (LE)

Cao et al., 2010

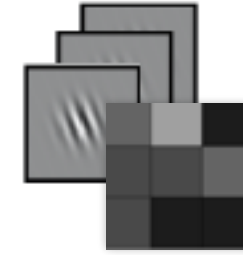


15

- Machine Learning: PCA, Normalization
- Landmark Detection zum Filtern
- Support Vector Machines zum Vergleich
- Ergebnisse zu Gabor & LBP bis zu 10% besser

Trends

- Kleine, unabhängige Filter
 - Flexibel, vereinfacht Training
 - Stabiler gegenüber Variationen der Pose



- Learning-basierte Deskriptoren
- Convolutional Neural Networks

16

- kleine Filter
 - Abstrakte Features
 - Invariant

Inhalt

- Motivation
- Ansätze
- „Identity-Preserving Face Space“
 - Architektur
 - Training
 - Auswertung
- Ausblick

Identity-Preserving Face Space

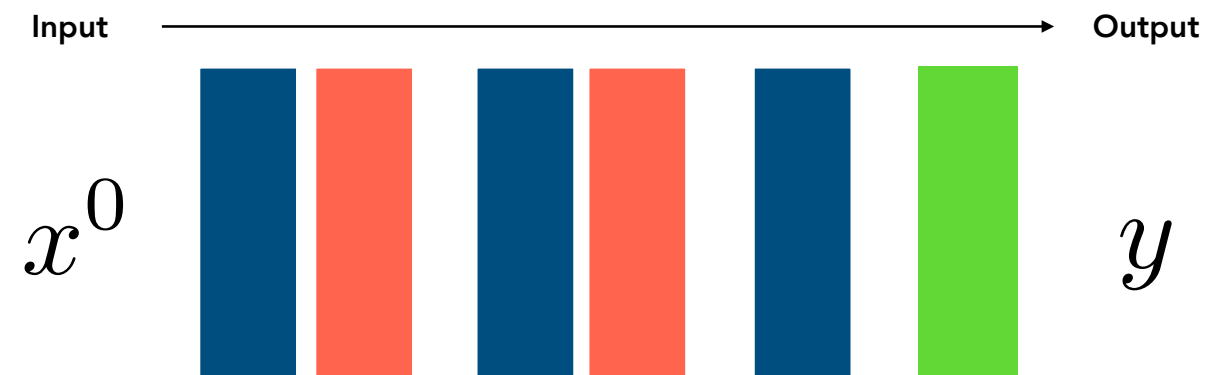
Zhu et al., 2013

- Gesichtsrepräsentation mittels Deep Neural Network
- Stabil über Varianzen hinweg
- Rekonstruktion des Gesichts

18

- 5 Jahre alt
- Repräsentation, nicht Vergleich
 - Notwendig für Vergleich
 - Merkmale im Gesicht enkodieren, (un)vergleichbar machen
 - Stabil über verschiedene Blickwinkel
 - Rekonstruktion in neutraler Ansicht

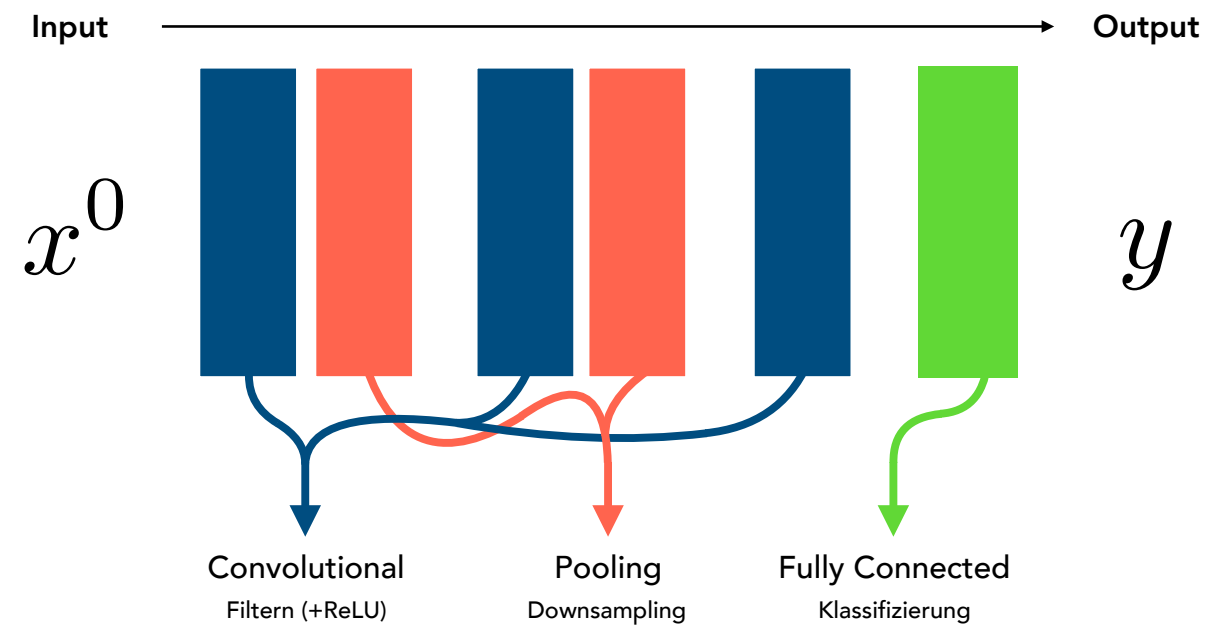
Architektur



19

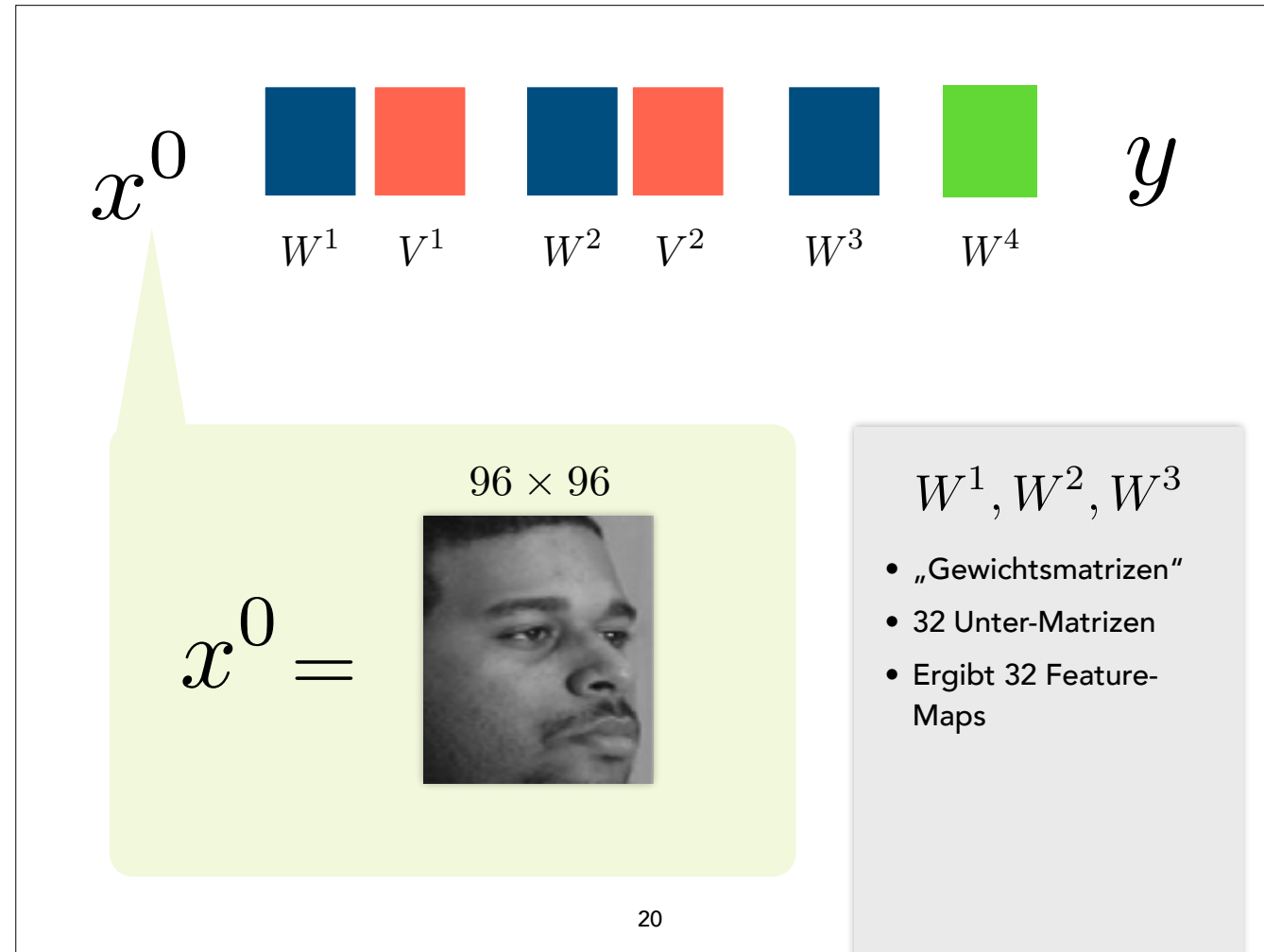
- Links nach Rechts
- x^0 Eingabe, y Ausgabe
- 6 Layer
 - Convolutional: Filtern von Texturen
 - Pooling: Downsampling, Reduzierung von Daten
 - Fully Connected: Klassifizierung
- Wie sieht es de facto aus?

Architektur

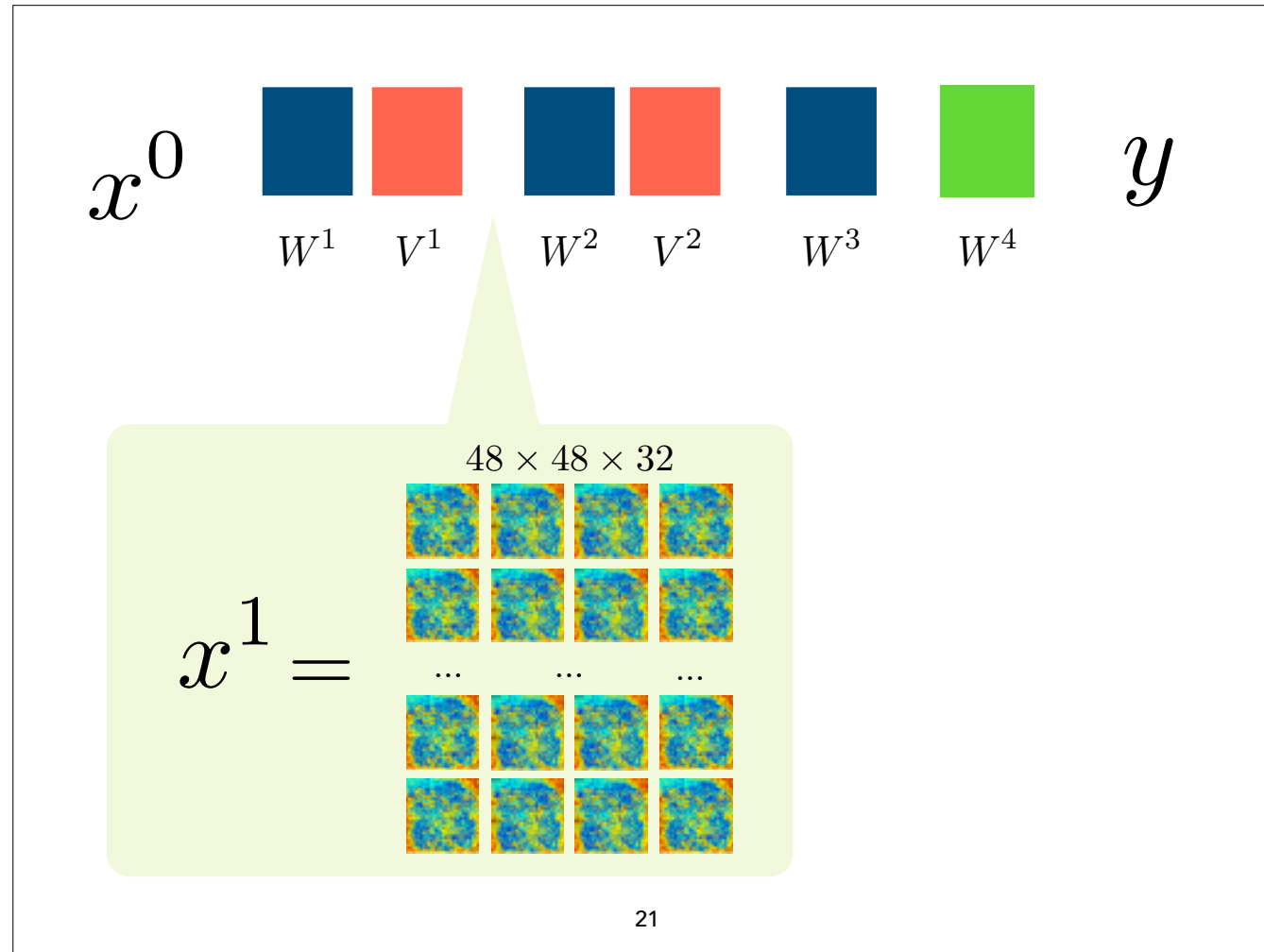


19

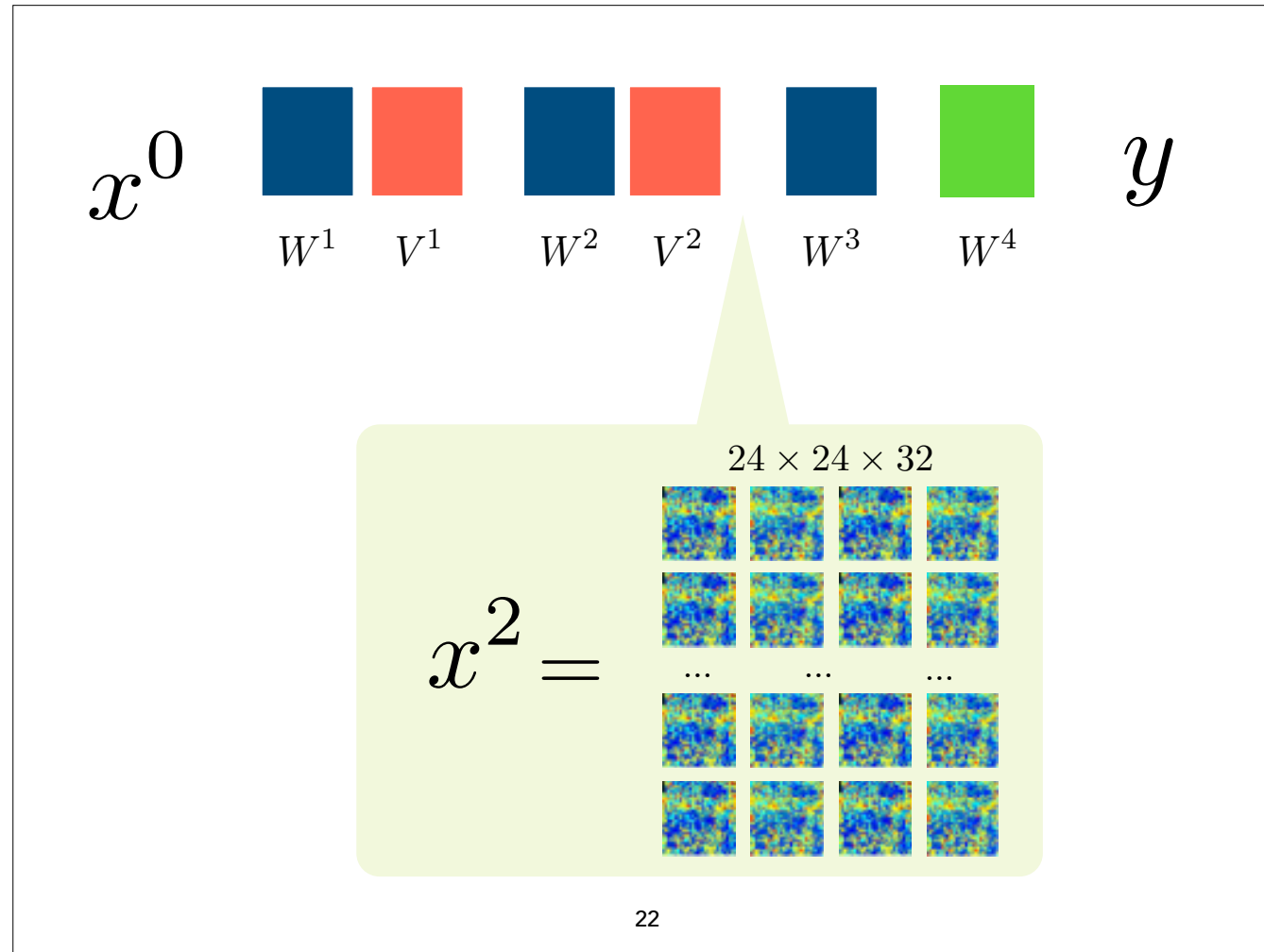
- Links nach Rechts
- x^0 Eingabe, y Ausgabe
- 6 Layer
 - Convolutional: Filtern von Texturen
 - Pooling: Downsampling, Reduzierung von Daten
 - Fully Connected: Klassifizierung
- Wie sieht es de facto aus?



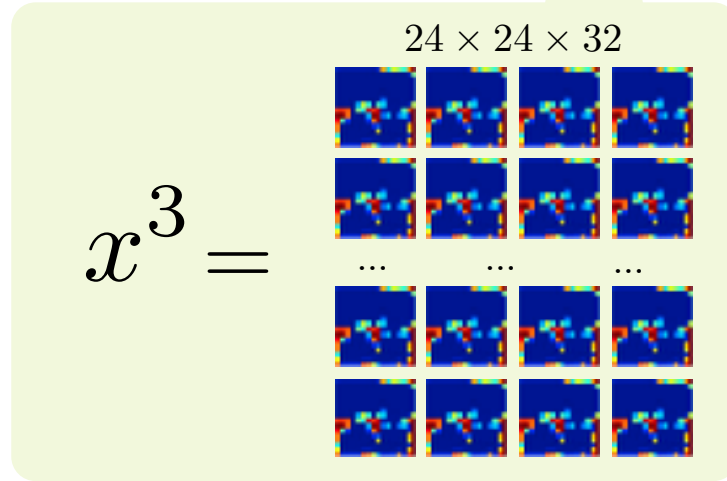
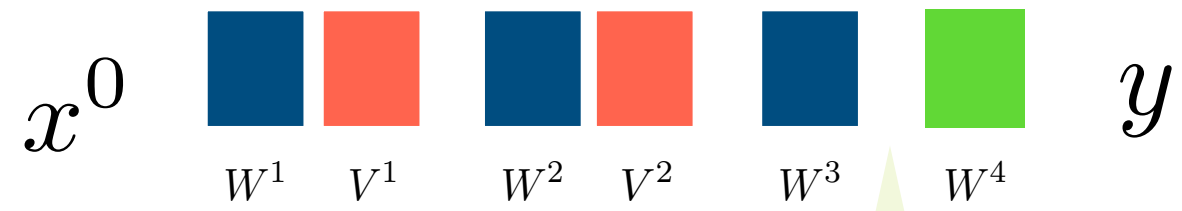
- Eingabe: 96 x 96 Pixel groß, s/w
- Gewichtsmatrizen der Convolutions
 - Jeweils 32 Unter-Matrizen
 - Dünnbesetzt
 - Keine geteilten Gewichte
 - Ergeben 32 Feature Maps

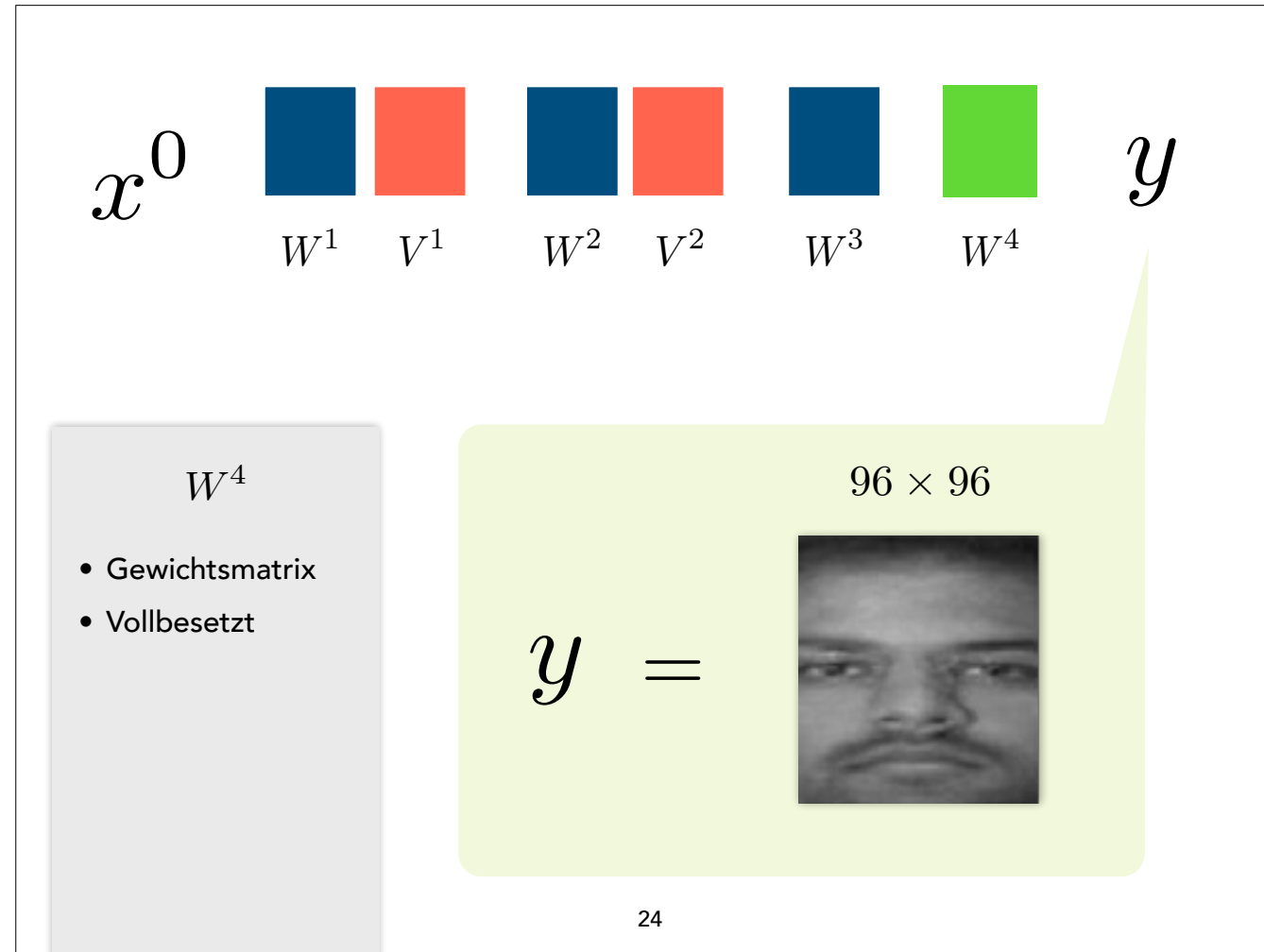


- Erste Convolution
- Erstes Downsampling
- Highlights für bestimmte Regionen
- 48 x 48



- Ähnlich wie W^1
- Kleinere Dimension, Features kristallisieren sich heraus

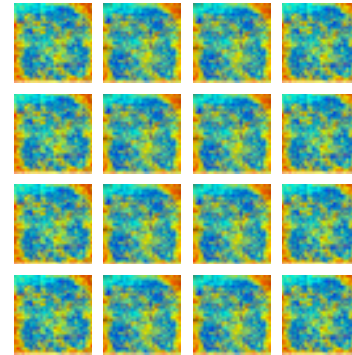




- Finally, Klassifikation und Rekonstruktion
- W^4 ist vollbesetzt, normales Layer
- Weist Pixel direkt zu

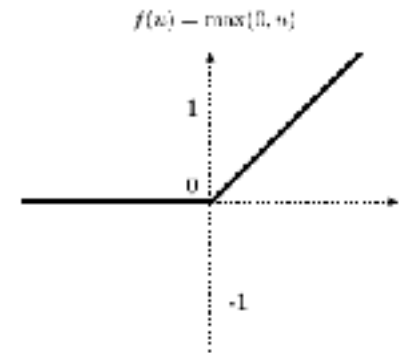
Invarianz

Ngiam et al., 2010



Un-tied Weight Matrices

Nair et al., 2010



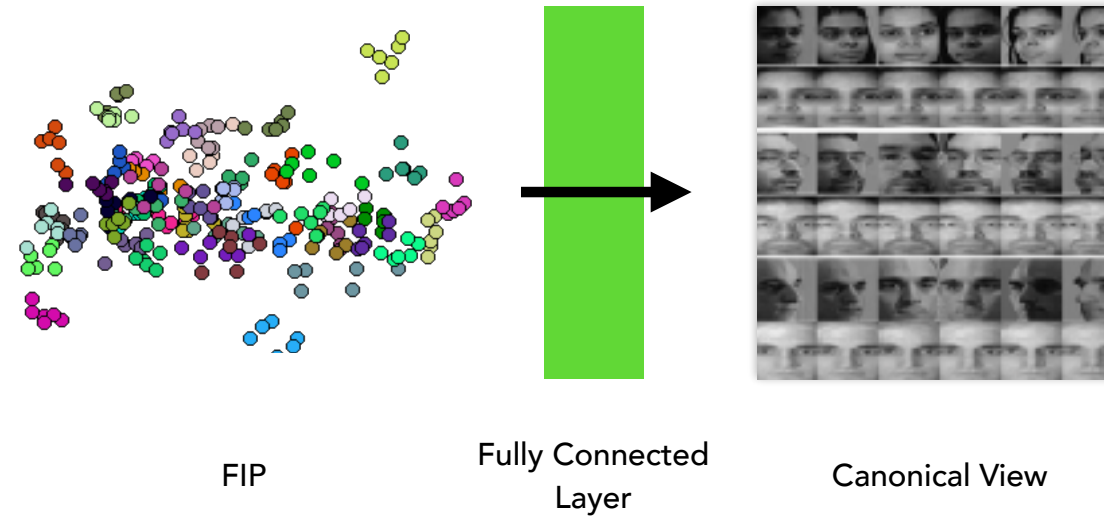
Rectified Linear Unit

25

- Schwierig zu visualisieren
- Wie können wir Invarianz für Features behaupten?
 - Individuelle Gewichtsmatrizen
 - Komplexe Merkmale, da mehr Gewichte
 - Pooling kombiniert sie
 - lokal verbunden, lokales Pooling
 - Tiled CNNs
 - ReLU relativiert Intensitäten: Schwierig, aber: Nonlinearity, Skalierung nicht möglich
 - hilfreich bei Belichtung

FIP & Canonical View

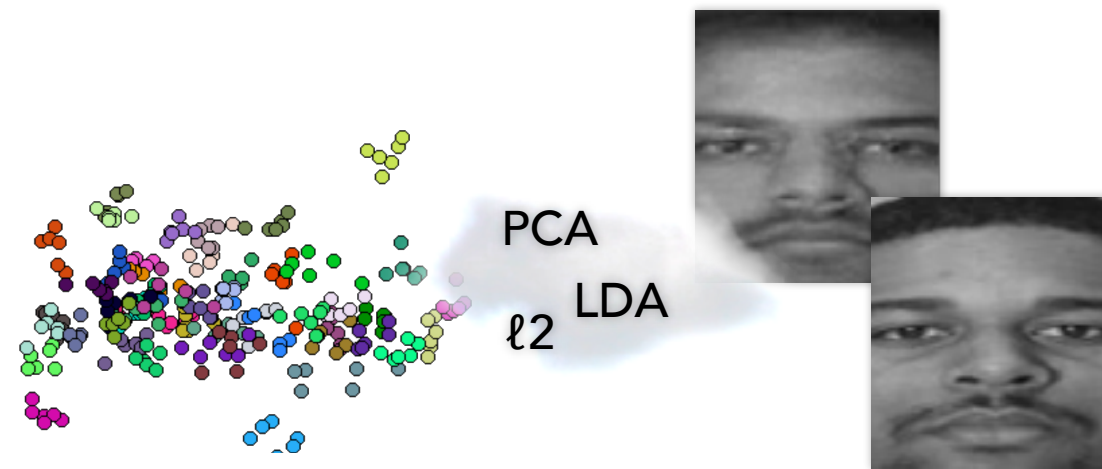
Zhu et al., 2013



26

- Was genau erhalten wir?
 - FIP: Datenmenge, beschreiben Intensitäten der gelernten Features
 - Fully Connected
 - Klassifikation
 - Canonical View entsteht daraus, quasi grafische Repräsentation

Vergleich



27

- Vergleich wird nicht vorgegeben
- Daten in mehrdimensionalem Format
 - Klassifikationen via
 - Euklidische Distanz (triviale Distanz)
 - PCA & LDA

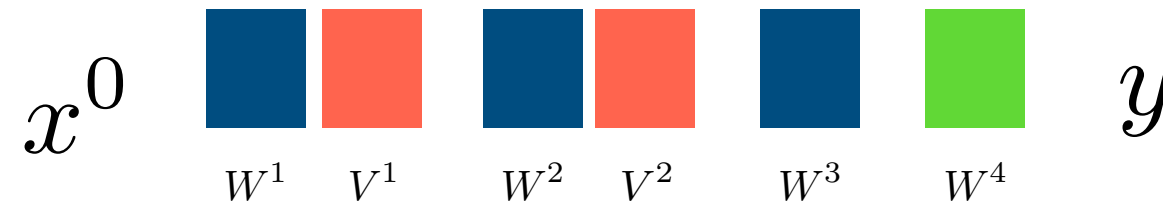
Inhalt

- Motivation
- Ansätze
- „Identity-Preserving Face Space“
 - Architektur
 - **Training**
 - Auswertung

28

- Bisher nur Aufbau beschrieben
- Wie werden diese Eigenschaften erreicht?
 - Invarianzen
 - Eigenschaften von FIP
- Nicht extrem tief, aber große Matrizen
- Aufteilen in Initialisierung und Update

Initialisierung

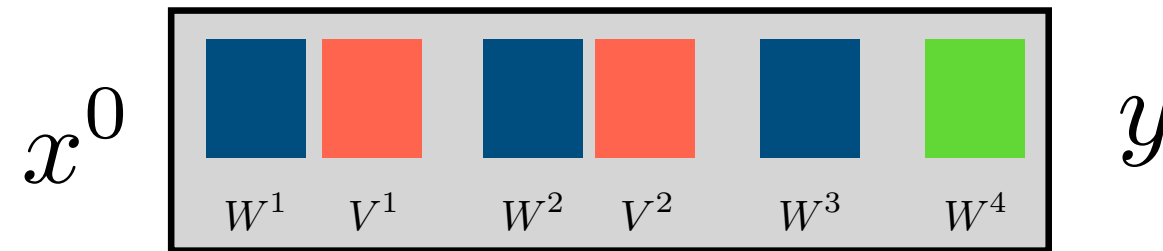


$$\arg \min_{W^1, W^2, W^3, W^4} \| \bar{Y} - \sigma(W^4 X^3) \|_F^2,$$

29

- Mathematik!
 - Normalerweise: Zufall (ineffizient)
- Frobenius-Norm: Wurzel aller summierten, quadrierten Einträge
 - Berechnung aller Gewichte ist nicht linear
 - Aber: W^i sind orthogonal
 - Daher Schrittweise: Lineares Ausgleichsproblem
 - QR-Zerlegung

Initialisierung



$$\arg \min_{W^1, W^2, W^3, W^4} \| \bar{Y} - \sigma(W^4 X^3) \|_F^2,$$

29

- Mathematik!
 - Normalerweise: Zufall (ineffizient)
- Frobenius-Norm: Wurzel aller summierten, quadrierten Einträge
 - Berechnung aller Gewichte ist nicht linear
 - Aber: W^i sind orthogonal
 - Daher Schrittweise: Lineares Ausgleichsproblem
 - QR-Zerlegung

Update: Backpropagation



30

- klassische Backpropagation
- Eingabe und Ground Truth bekannt
 - Feedforward: Ausgabe y berechnen, fehlerbehaftet, da Matrizen schlecht
 - Error Estimation: E berechnen
 - Backprop: Fehler propagieren und mittels partieller Ableitung Δw berechnen

Update: Backpropagation



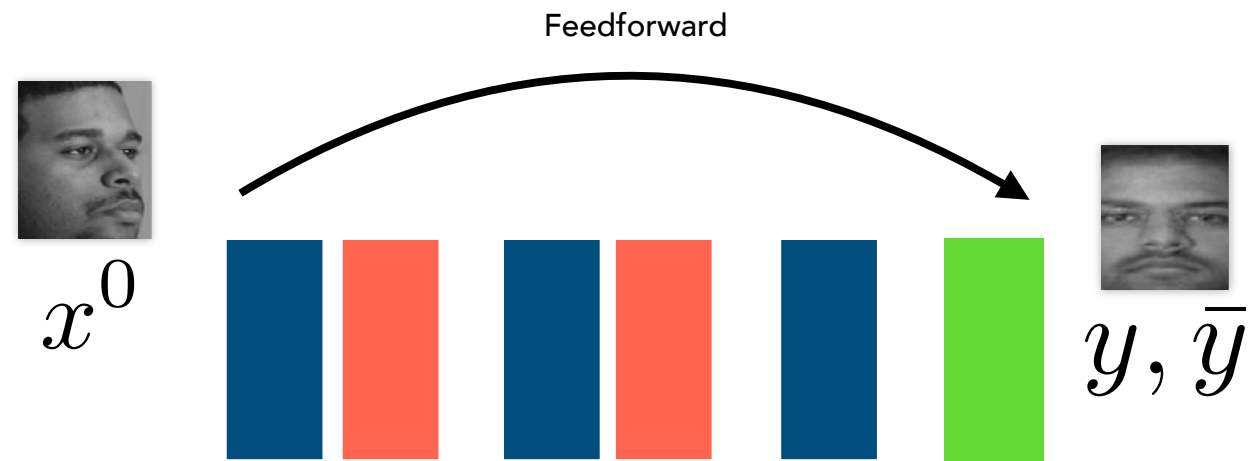
x^0



30

- klassische Backpropagation
- Eingabe und Ground Truth bekannt
 - Feedforward: Ausgabe y berechnen, fehlerbehaftet, da Matrizen schlecht
 - Error Estimation: E berechnen
 - Backprop: Fehler propagieren und mittels partieller Ableitung Δ - w berechnen

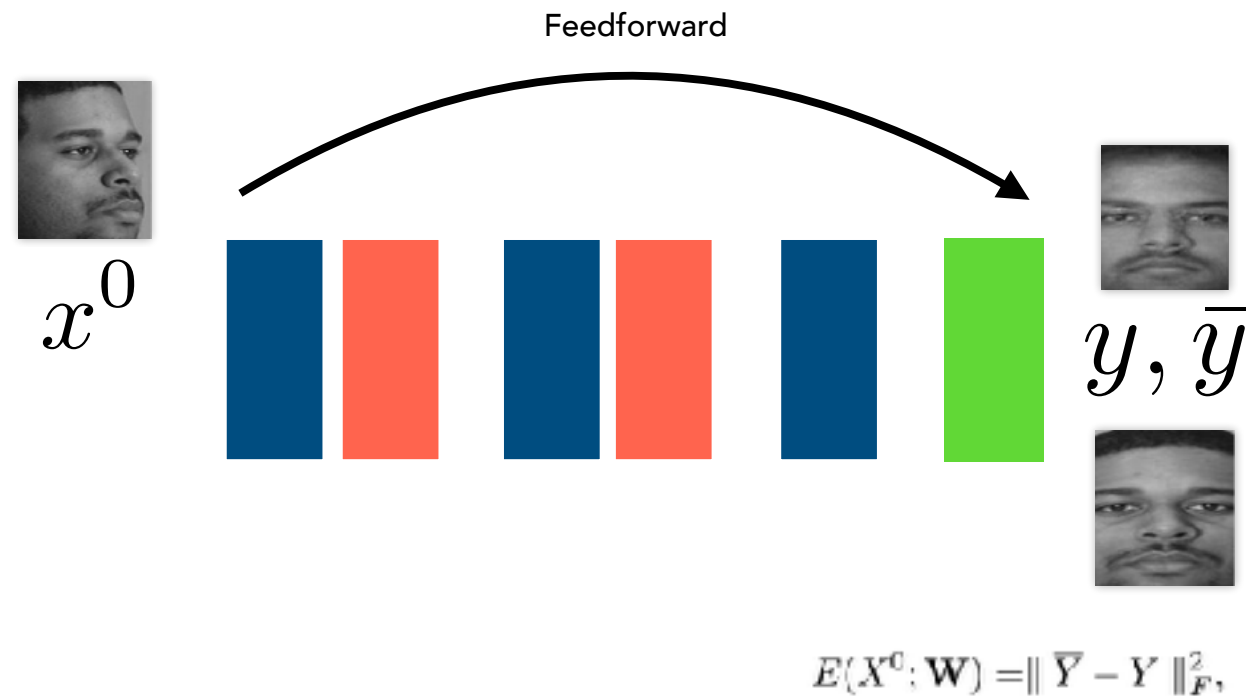
Update: Backpropagation



30

- klassische Backpropagation
- Eingabe und Ground Truth bekannt
 - Feedforward: Ausgabe y berechnen, fehlerbehaftet, da Matrizen schlecht
 - Error Estimation: E berechnen
 - Backprop: Fehler propagieren und mittels partieller Ableitung Δ - w berechnen

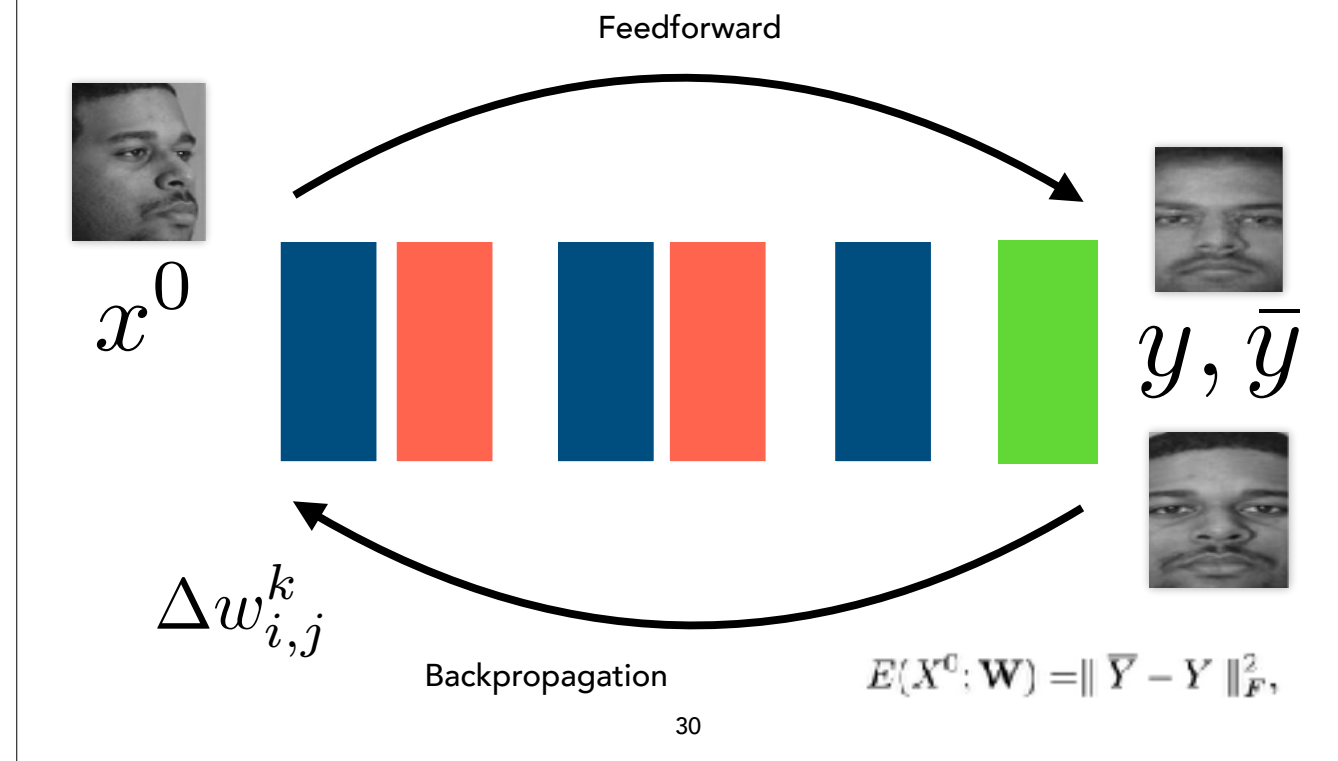
Update: Backpropagation



30

- klassische Backpropagation
- Eingabe und Ground Truth bekannt
 - Feedforward: Ausgabe y berechnen, fehlerbehaftet, da Matrizen schlecht
 - Error Estimation: E berechnen
 - Backprop: Fehler propagieren und mittels partieller Ableitung Δ -w berechnen

Update: Backpropagation



- klassische Backpropagation
- Eingabe und Ground Truth bekannt
 - Feedforward: Ausgabe y berechnen, fehlerbehaftet, da Matrizen schlecht
 - Error Estimation: E berechnen
 - Backprop: Fehler propagieren und mittels partieller Ableitung Δw berechnen

Inhalt

- Motivation
- Ansätze
- „Identity-Preserving Face Space“
 - Architektur
 - Training
 - **Auswertung**

31

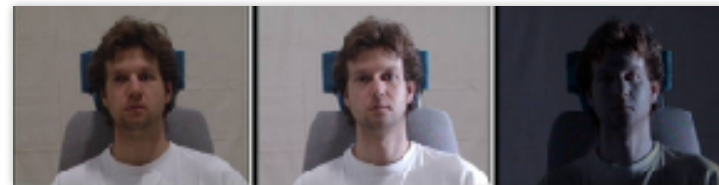
- Wie verhält es sich im Vergleich?
- Wir vermuten, Learning-based ist besser (LE)
 - Ist es so?

MultiPIE

Gross et al., 2010



Blickwinkel

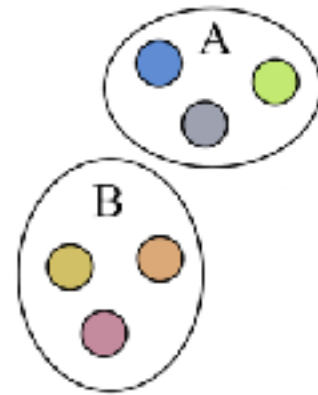


Belichtungen

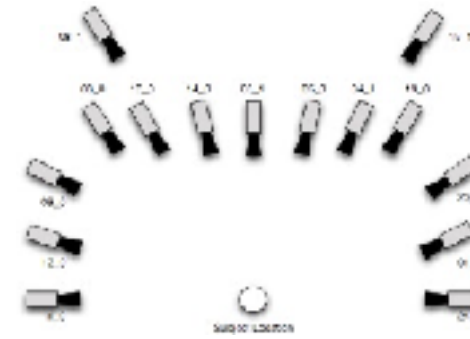
32

- MultiPIE: Pose, Illumination, Expression
- 337 Identitäten, knapp 750.000 Bilder
 - Verschiedene Blickwinkel, Belichtungen
- De facto-Standard für Gesichtserkennung
 - Werte für Gabor, LBP, ...

Experiment



Training vs. Proben

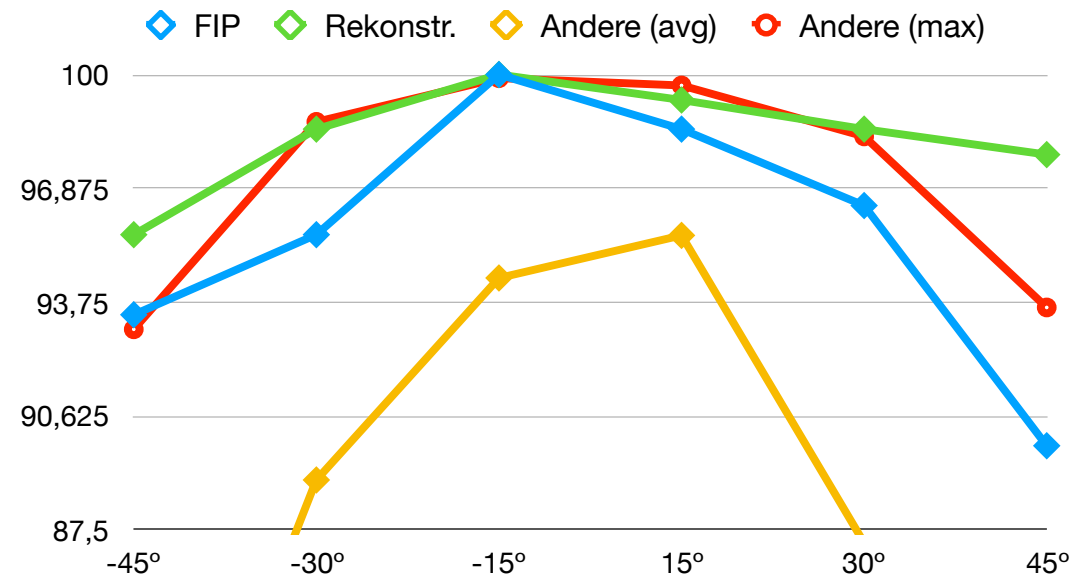


Blickwinkel

- Training mit knapp 50% der Identitäten, Proben aus dem Rest
- Training mit allen Blickwinkeln, Proben ohne Frontal
- Mehrere Experimente...

Ergebnisse

Blickwinkel



34

- Ergebnisse nur für ersten Durchgang, da die meisten Gegner
 - Nur Blickwinkel
- Andere, das sind
 - CRBM (Convolutional Restricted Boltzmann Machines)
 - EGFC (Ensemble Gabor Fisher Classifier) > 3D
 - LGBP (Local Gabor Binary Patterns)
 - LE (Learning-based)
- Performed ähnlich gut wie 3D, wenn nicht sogar besser

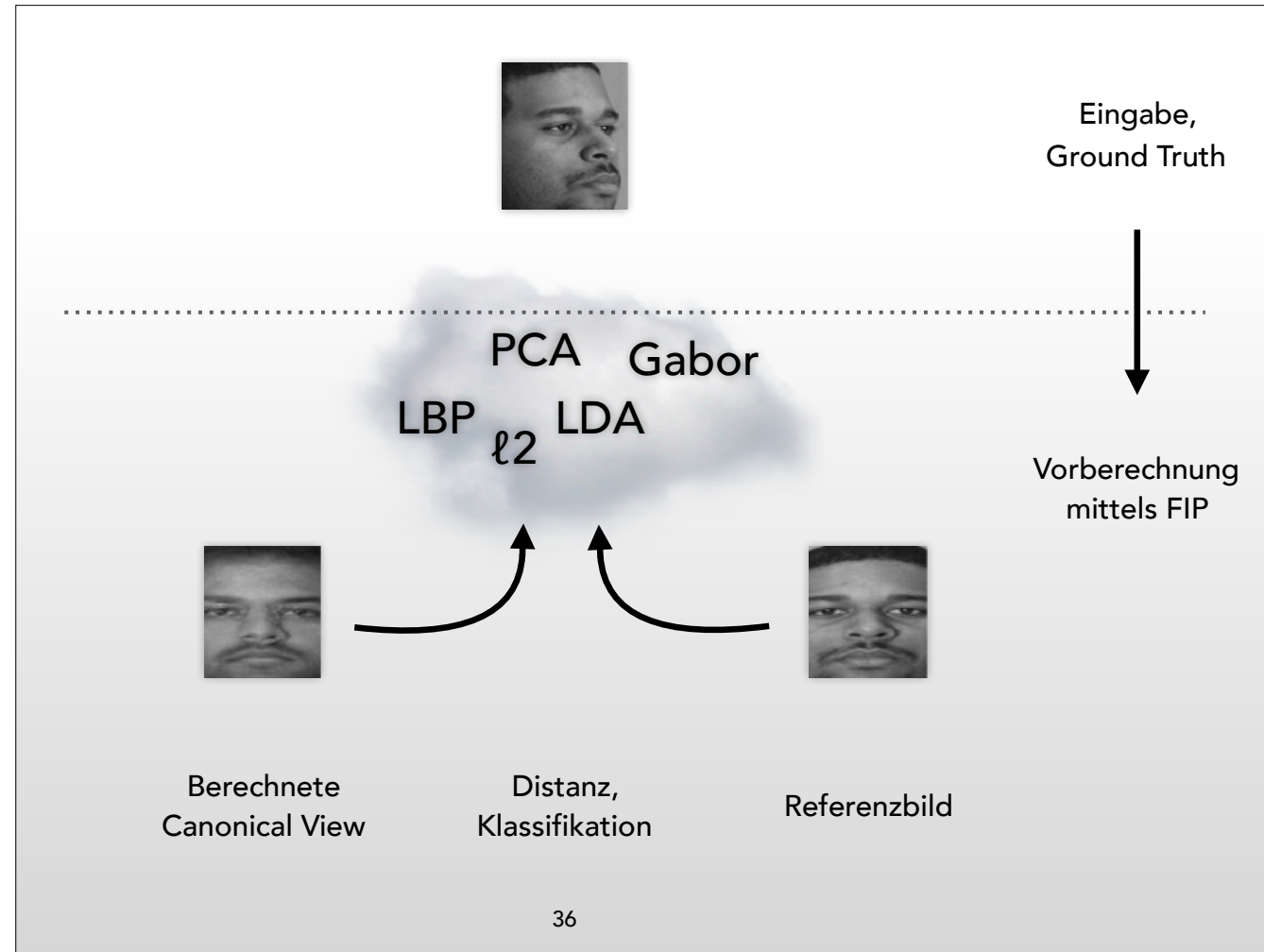
Canonical View

Zhu et al., 2013



35

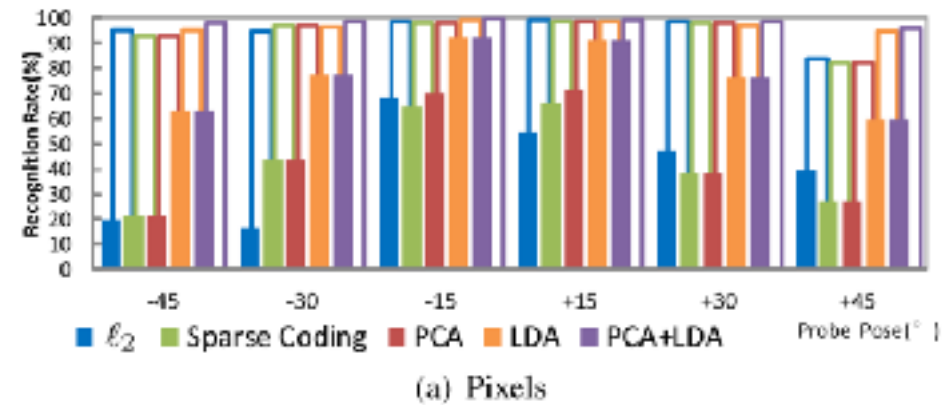
- Was kann die Canonical View noch?
 - Wir wissen, Ergebnisse sind gut
 - (sieht man hier)
- Vorteil: Beständig, invariant
- Zweidimensionale Repräsentation, wie sie viele Systeme erwarten



- Weiteres Experiment
- Bestehende Algorithmen mit Canonical View verwendet
 - PCA, LDA, euklidische Distanz
 - Aber auch Deskriptoren (Gabor, Local Binary Patterns)
 - Performance liegt prinzipiell weit dahinter

Canonical View

Zhu et al., 2013



37

- Ausschnitt der Ergebnisse
- „Vorderen“ Balken sind die Ergebnisse auf normalen Bildern
- „leeren“ die mit denselben Bildern
 - wurden mit Canonical View berechnet
 - liegen immer zwischen 80 und sogar fast 100%
- Wer andere bzw. schmalere Darstellung bevorzugt
 - ... kann auf alte Deskriptoren zurückgreifen

Zusammenfassung

- **Gesichtserkennung:**
 - Alltagsanwendung
 - Problematik: Varianzen
- **Ansätze:**
 - Gabor-Filter
 - Local Binary Patterns
 - Learning-basierte Deskriptoren (LE)
- **Trends:**
 - Deep Learning
 - Kleine, flexible Filter (Convolutional NNs)
- **Zhu et al., 2013:**
 - Descriptor-Learning mittels Deep Learning
 - Convolutional Neural Network on Steroids
 - Training mittels Backpropagation und MultiPIE-Datensatz
 - Features stabil gegenüber Belichtung & Pose
 - Unterstützt bestehende Methoden
 - Rekonstruktion von Gesichtern in neutraler Pose

38

- Problematiken der Gesichtserkennung: Varianzen
- Ansätze
 - Traditionelle Deskriptoren
 - Learning-basierte Deskriptoren
- Trends: Deep Learning, Convolutions
- Zhu et al.

Zusammenfassung



- **Gesichtserkennung:**
 - Alltagsanwendung
 - Problematik: Varianzen
- **Ansätze:**
 - Gabor-Filter
 - Local Binary Patterns
 - Learning-basierte Deskriptoren (LE)
- **Trends:**
 - Deep Learning
 - Kleine, flexible Filter (Convolutional NNs)
- **Zhu et al., 2013:**
 - Descriptor-Learning mittels Deep Learning
 - Convolutional Neural Network on Steroids
 - Training mittels Backpropagation und MultiPIE-Datensatz
 - Features stabil gegenüber Belichtung & Pose
 - Unterstützt bestehende Methoden
 - Rekonstruktion von Gesichtern in neutraler Pose



38

- Problematiken der Gesichtserkennung: Varianzen
- Ansätze
 - Traditionelle Deskriptoren
 - Learning-basierte Deskriptoren
- Trends: Deep Learning, Convolutions
- Zhu et al.

Literatur

- Zhu, Z., Luo, P., Wang, X., & Tang, X. (2013). Deep learning identity-preserving face space. In Proceedings of the IEEE International Conference on Computer Vision (pp. 113-120).
- Zhu, Z., Luo, P., Wang, X., & Tang, X. (2014). Recover canonical-view faces in the wild with deep neural networks. arXiv preprint arXiv:1404.3543.
- Wiskott, L., Krüger, N., Kuiger, N., & Von Der Malsburg, C. (1997). Face recognition by elastic bunch graph matching. IEEE Transactions on pattern analysis and machine intelligence, 19(7), 775-779.
- Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. IEEE transactions on pattern analysis and machine intelligence, 28(12), 2037-2041.
- Cao, Z., Yin, Q., Tang, X., & Sun, J. (2010, June). Face recognition with learning-based descriptor. In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on (pp. 2707-2714). IEEE.
- Huang, G. B., Lee, H., & Learned-Miller, E. (2012, June). Learning hierarchical representations for face verification with convolutional deep belief networks. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on (pp. 2518-2525). IEEE.
- Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. (2010). Multi-pie. Image and Vision Computing, 28(5), 807-813.
- Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., & Ma, Y. (2009). Robust face recognition via sparse representation. IEEE transactions on pattern analysis and machine intelligence, 31(2), 210-227.
- Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (pp. 315-323).
- Ngiam, J., Chen, Z., Chia, D., Koh, P. W., Le, Q. V., & Ng, A. Y. (2010). Tiled convolutional neural networks. In Advances in neural information processing systems (pp. 1279-1287).

Medienquellen

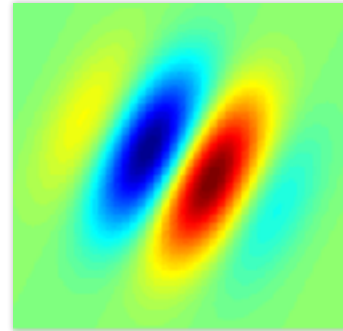
- Folie 4: Gesichter (Zhu et al., 2013)
- Folie 6: iPhone X (https://support.apple.com/kb/SP770?locale=de_DE), Überwachungskamera (<http://www.dlink.com/de/de/business-solutions/ip-surveillance/all-products>), Amazon Go (<https://arstechnica.com/information-technology/2018/01/amazons-cashier-free-convenience-store-in-seattle-opens-to-the-public-today/>)
- Folie 9: Gesichter (<http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Content.html>)
- Folie 11: Gabor-Filter, 3D Scan (<https://www.cgtrader.com/3d-models/character/woman/high-detailed-female-face-with-uv-and-low-poly>)
- Folie 12: Gabor-Filter (https://en.wikipedia.org/w/index.php?title=Gabor_filter&oldid=813455572), Gesicht mit Bunch Graph (Wiskott et al., 1997)
- Folie 13: LBP (Ahonen et al., 2006)
- Folie 14: Histogramm (<https://stackoverflow.com/questions/23907929/matlab-create-3d-histogram-from-sampled-data>)
- Folie 15: LE-Deskriptor (Cao et al., 2010),
- Folie 16: Gabor-Wavelets (Wiskott et al., 1997), LBP (Ahonen et al., 2006)
- Folien 20-24: Gesichter & CNN-Werte (Zhu et al., 2013)
- Folie 25: CNN-Werte (Zhu et al., 2013), ReLU (https://www.researchgate.net/figure/ReLU-activation-function_fig3_319235847)
- Folie 26: Gesichter & Ausgabewerte (Zhu et al., 2013)
- Folie 27: S. Folie 26
- Folie 31: MultiPIE-Gesichter (<http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Content.html>)
- Folie 32: ebd.
- Folie 33: MultiPIE-Ansicht (<http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Content.html>)
- Folie 35: Gesichter (Zhu et al., 2013)
- Folie 36: Gesichter (Zhu et al., 2013)
- Folie 37: Ergebnisse Deskriptoren mit Canonical View (Zhu et al., 2013)
- Folie 38: Überwachungskamera (<https://eu.dlink.com/be/nl/products/dcs-7513-outdoor-full-hd-wdr-poe-day-night-fixed-bullet-camera-with-ir-led>)

Appendix

- Gabor-Filter
- Local Binary Pattern Descriptor
- Lineare Diskriminanzanalyse (LDA)
- Principal Component Analysis (PCA)
- Euklidischer Abstand (ℓ_2)
- Rectified Linear Units (ReLUs)
- Tiled CNNs
- Backpropagation
- MultiPIE
- Ergebnisse Belichtung (Setting-II)
- Convolutional Neural Networks

Gabor-Filter

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \psi\right)$$

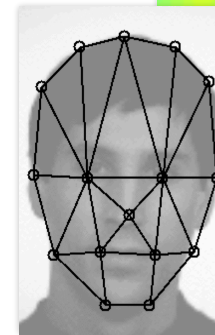


2D-Filter

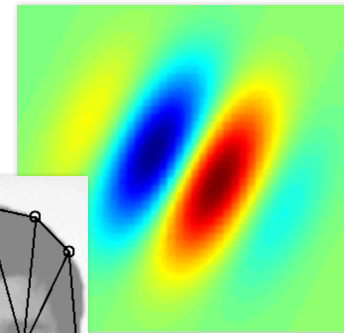
Gabor Descriptor

Wiskott et al., 1997

- Feature-Erkennung mittels Gabor-Filtern, Wavelets
 - Auf Ausschnitten angewandt, gesammelt in „Jets“
- Assoziation auf Gesichter mittels „Bunch Graphs“



Bunch Graph

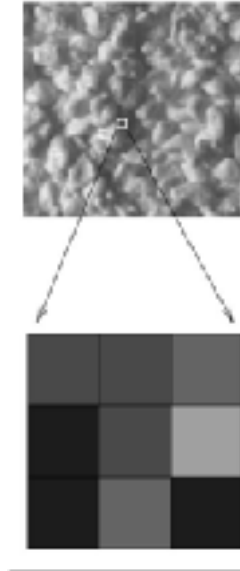


Gabor-Filter

Local Binary Pattern Pattern Descriptor

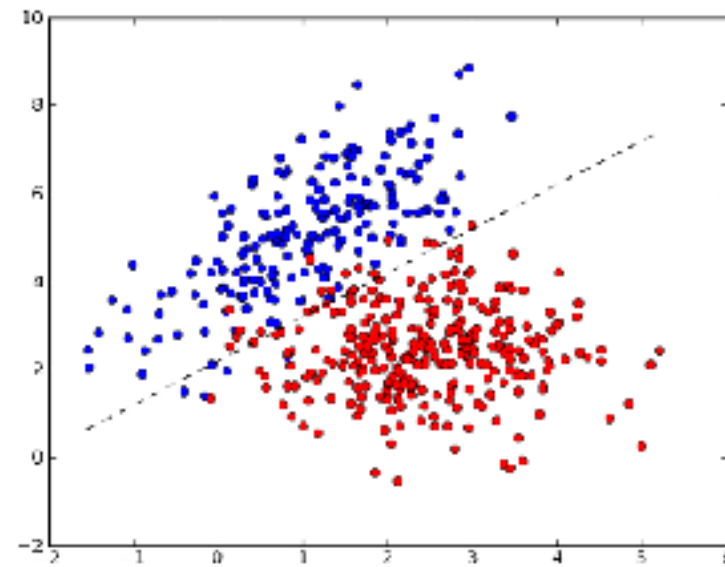
Ahonen et al., 2006

- Local Binary Pattern: Texturanalyse
- Filter mit kleinem Kernel
 - „composition of micropatterns“
- Encodieren der Filter-Werte als Histogramm
 - Feature-Deskriptor



LDA

Lineare Diskriminanzanalyse

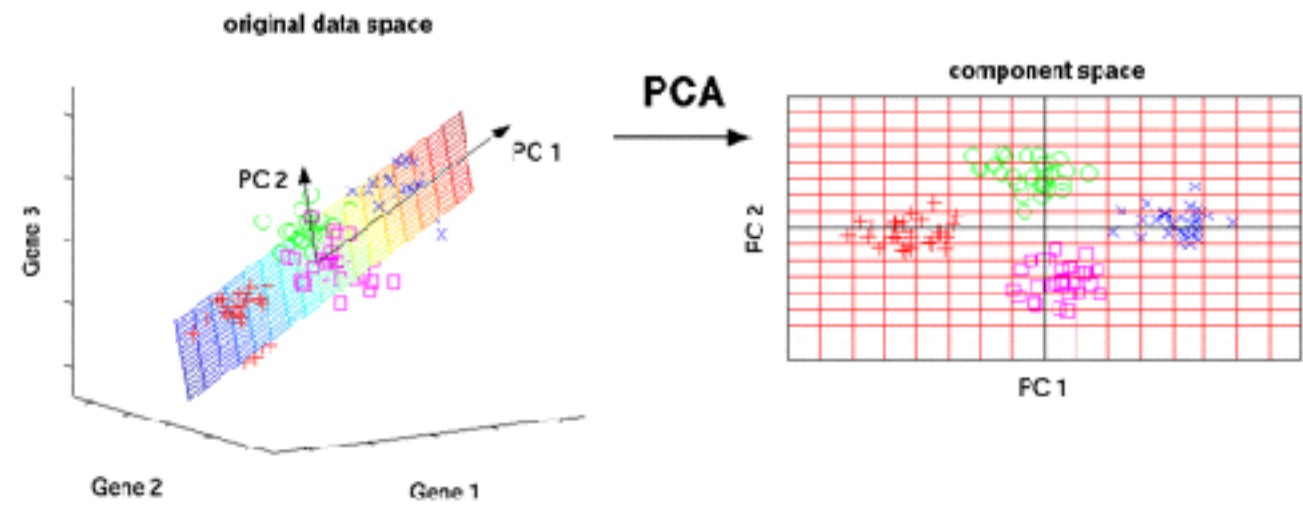


45

- Klassifikation bzw.
 - Trennung von Daten anhand linearer Ebene
 - Nicht unähnlich zu SVMs

PCA

Principal Component Analysis



46

- Reduzierung der Dimensionen
- Projektion der Datenwerte

ℓ_2

Euklidische Distanz

$$\sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

ReLU

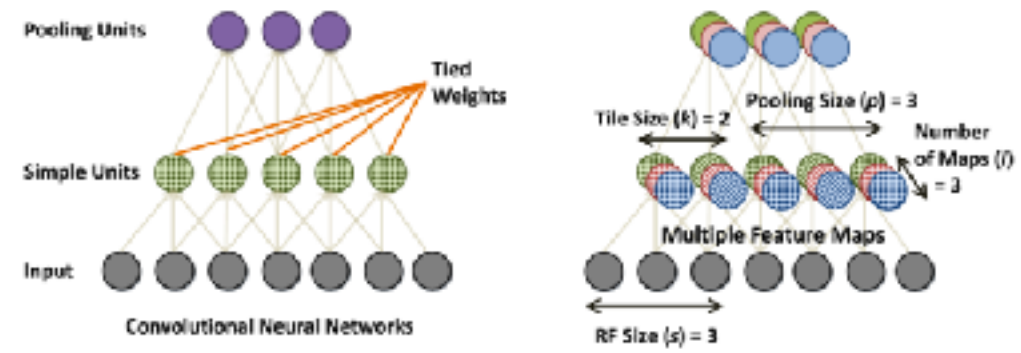
Glorot et al., 2011

$$\sigma(x) = \max(0, x)$$

$$\max(0, ax) = a \max(0, x)$$

Tiled CNNs

Ngiam et al., 2010



Backpropagation

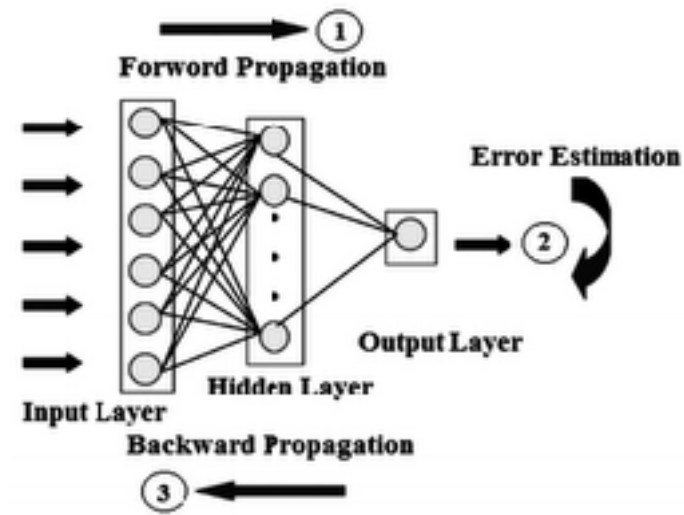
Quadratischer Fehler: $E = \frac{1}{2}(t - y)^2$

Partielle Ableitung
über Kettenregel:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial \text{net}_j} \frac{\partial \text{net}_j}{\partial w_{ij}}$$
$$= o_i \delta_j$$

$$\delta_j = \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial \text{net}_j} = \begin{cases} (o_j - t_j) o_j (1 - o_j) & \text{if } j \text{ is an output neuron,} \\ (\sum_{\ell \in L} \delta_\ell w_{j\ell}) o_j (1 - o_j) & \text{if } j \text{ is an inner neuron.} \end{cases}$$

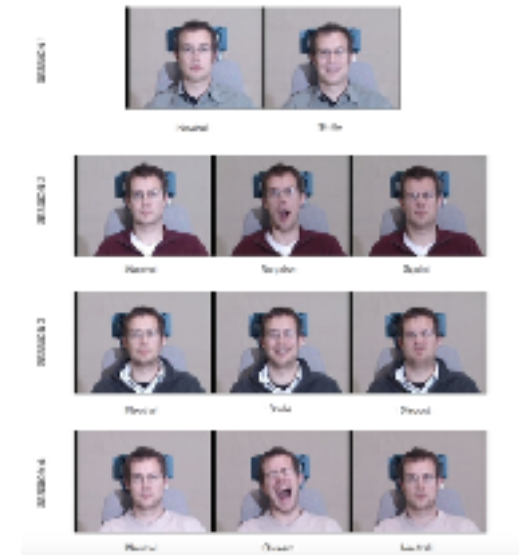
Backpropagation



MultiPIE

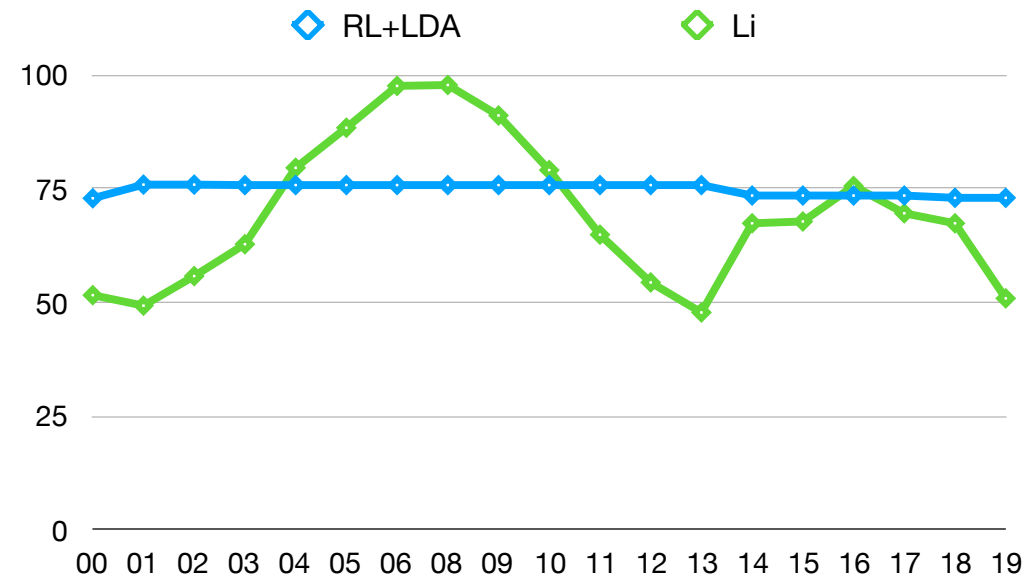
Gross et al., 2010

- 754.204 Fotos von 337 Personen
- Aufgenommen in vier verschiedenen Sitzungen
- 15 Betrachtungswinkel (-45° bis 45°)
- 20 Belichtungsgrade



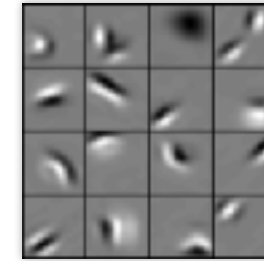
Ergebnisse

Belichtung



Convolutional Neural Networks

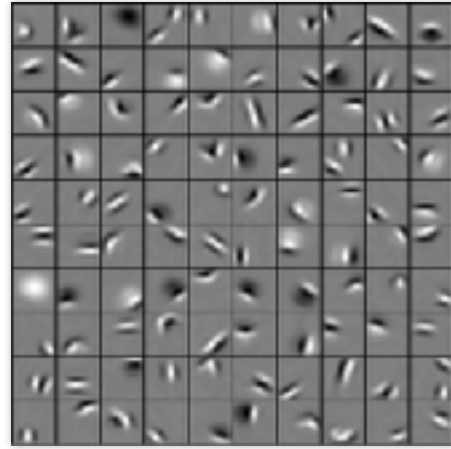
- Ansatz: Verwendung von Faltungs-Matrizen (Convolution Filter) zur Feature-Erkennung
 - Matrizen sind klein bzw. dünnbesetzt
- Pooling-Layer performen Downsampling
 - Dimensionen reduzieren, Features erhalten



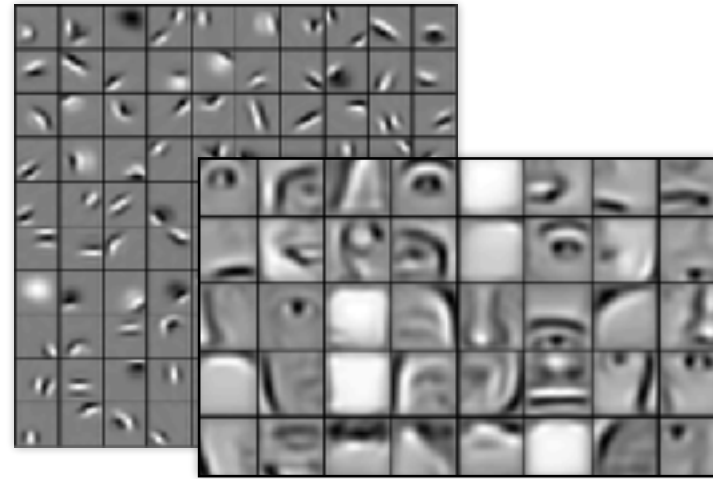
Convolution
Filter

Convolutional Neural Networks

Convolutional Neural Networks



Convolutional Neural Networks



Convolutional Neural Networks

