DEEP LEARNING GRAZ MEETUP, 4.7.2018, GRAZ, FGM, HERBERT DANZINGER

WORUM ES GEHT

- Technische Aspekte der Personenidentifikation mit Hilfe von Convolutional Networks am Beispiel von Amazon Rekognition
- CNNS
  - convolutions
  - pooling
- theoretische Einführung – praktischer Teil im Herbst 2018

WORUM GEHT ES NICHT

- Codebeispiele .. aber es gibt ein paar Links am Ende...
- Predictive policing
- Verbrechen verhindern

Amazon ermöglicht Live-Gesichtserkennung für ganze Städte

23.05.2018 08:37 Uhr – Daniel AJ Sokolov

vorlesen

(Bild: Kurt Bauschardt CC BY-SA 2.0)

HTTPS://WWW.HEISE.DE/NEWSTICKER/MELDUNG/AMAZON-ERMOEGLICHT-LIVE-GESICHTSERKENNUNG-FUER-GANZE-STAEDTE-4055143.HTML

AMAZON REKOGNITION IN DEN MEDIEN

ermöglicht, ganze Städte in Echtzeit zu überwachen.

in den USA wird Rekognition bereits eingesetzt

US-Bürgerrechtler (ACLU) fordern ein Ende dieser Überwachung.

ACLU

Email address ZIP code GET UPDATES

BECOME A MEMBER / RENEW / TAKE ACTION / DONATE

ISSUES KNOW YOUR RIGHTS DEFENDING OUR RIGHTS BLOGS ABOUT SHOP

Amazon Teams Up With Government to Deploy Dangerous New Facial Recognition Technology


Matt Cagle, Technology and Civil Liberties attorney, ACLU of Northern California
& Nicole Ozer, Technology & Civil Liberties Director, ACLU of California

MAY 22, 2018 | 10:00 AM

TAGS: Face Recognition Technology, Surveillance Technologies, Privacy & Technology

HTTPS://WWW.ACLU.ORG/BLOG/PRIVACY-TECHNOLOGY/SURVEILLANCE-TECHNOLOGIES/AMAZON-TEAMS-GOVERNMENT-DEPLOY-DANGEROUS-NEW

FEATURES VON AMAZON REKOGNITION

HTTPS://AWS.AMAZON.COM/DE/BLOGS/MACHINE-LEARNING/AMAZON-REKOGNITION-ANNOUNCES-REAL-TIME-FACE-RECOGNITION-SUPPORT-FOR-RECOGNITION-OF-TEXT-IN-IMAGE-AND-IMPROVED-FACE-DETECTION/

TEXT IN IMAGE

HTTPS://AWS.AMAZON.COM/DE/BLOGS/MACHINE-LEARNING/AMAZON-REKOGNITION-ANNOUNCES-REAL-TIME-FACE-RECOGNITION-SUPPORT-FOR-RECOGNITION-OF-TEXT-IN-IMAGE-AND-IMPROVED-FACE-DETECTION/

Real-time Face Recognition

You can now perform real-time face searches against collections with tens of millions of faces. This represents a 5-10X reduction in search latency, while simultaneously allowing for collections that can store 10-20X more faces than before. For security and public safety applications, this update helps identify persons of interest against a collection of millions of faces in real-time, enabling use cases that require an immediate response.

The Washington County Sheriff's Office is the primary first responder for 911 calls from the citizens of Oregon. The office also provides support for crime prevention to other city police departments countywide. The Sheriff's Office has been using Amazon Rekognition over the past year to reduce the identification time of reported suspects from 2-3 days down to minutes and had apprehended their first suspect within a week by using their new system.

"These improvements allow deputies in the field to receive the response to searches in near real time. This allows them to get the information they need and take action quickly. Seconds saved in the field can make the difference in saving a life." – *Chris Adzima, Senior Systems Analyst for the Washington County Sheriff's Office.*

REAL TIME FACE RECOGNITION

HTTPS://AWS.AMAZON.COM/DE/BLOGS/MACHINE-LEARNING/AMAZON-REKOGNITION-ANNOUNCES-REAL-TIME-FACE-RECOGNITION-SUPPORT-FOR-RECOGNITION-OF-TEXT-IN-IMAGE-AND-IMPROVED-FACE-DETECTION/

CROWD MODE FACE DETECTION

- 100 Gesichter
- klassifiziert demographische Features (Alter, Geschlecht...)
- klassifiziert Stimmung

HTTPS://AWS.AMAZON.COM/DE/BLOGS/MACHINE-LEARNING/AMAZON-REKOGNITION-ANNOUNCES-REAL-TIME-FACE-RECOGNITION-SUPPORT-FOR-RECOGNITION-OF-TEXT-IN-IMAGE-AND-IMPROVED-FACE-DETECTION/

PERSON TRACKING

AWS Documentation » Amazon Rekognition » Developer Guide » Tracking People

Tracking People

Amazon Rekognition Video can track people in videos and provide information such as:

- The location of the person in the video frame at the time they were tracked.
- Facial landmarks such as the position of the left eye, when detected.

Amazon Rekognition Video person tracking in stored videos is an asynchronous operation. To start the tracking of people in videos call StartPersonTracking. Amazon Rekognition Video publishes the completion status of the video analysis to an Amazon Simple Notification Service topic. If the video analysis is succesful, call GetPersonTracking to get results of the video analysis. For more information about calling Amazon Rekognition Video API operations, see Calling Amazon Rekognition Video Operations.

The following procedure shows how to track people through a video stored in an Amazon S3 bucket. The example expands on the code in Analyzing a Video Stored in an Amazon S3 Bucket with Java or Python (SDK) which uses an Amazon Simple Queue Service queue to get the completion status of a video analysis request.

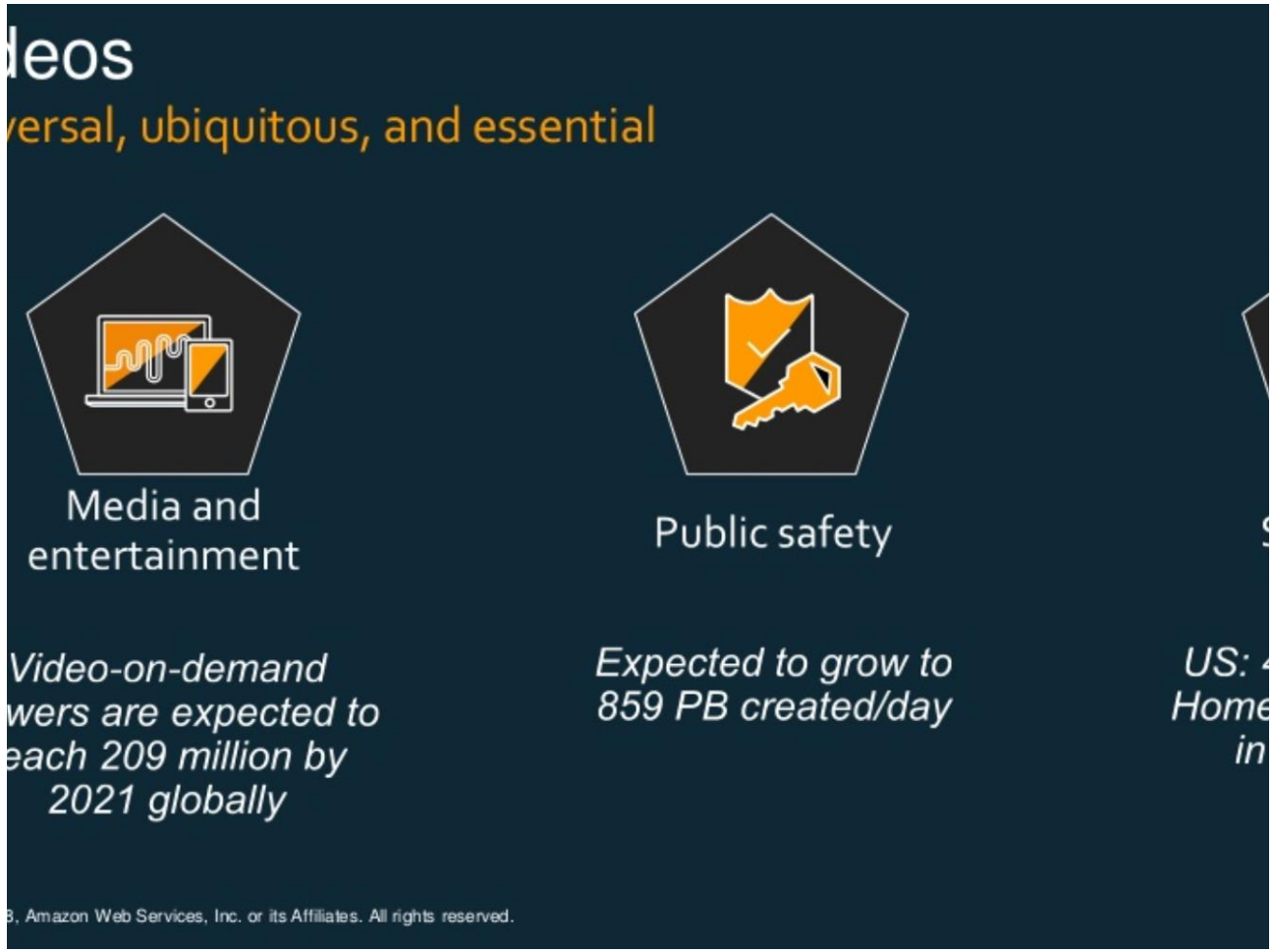
HTTPS://DOCS.AWS.AMAZON.COM/REKOGNITION/LATEST/DG/PERSONS.HTML

PERSON TRACKING CODE SAMPLE

```
def GetResultsPersons(self, jobId):
    maxResults = 10
    paginationToken = ''
    finished = False

    while finished == False:
        response = self.rek.get_person_tracking(JobId=jobId,
                                                MaxResults=maxResults,
                                                NextToken=paginationToken)

        print(response['VideoMetadata']['Codec'])
        print(str(response['VideoMetadata']['DurationMillis']))
        print(response['VideoMetadata']['Format'])
        print(response['VideoMetadata']['FrameRate'])

        for personDetection in response['Persons']:
            print('Index: ' + str(personDetection['Person']['Index']))
            print('Timestamp: ' + str(personDetection['Timestamp']))
            print()

        if 'NextToken' in response:
            paginationToken = response['NextToken']
        else:
            finished = True
```

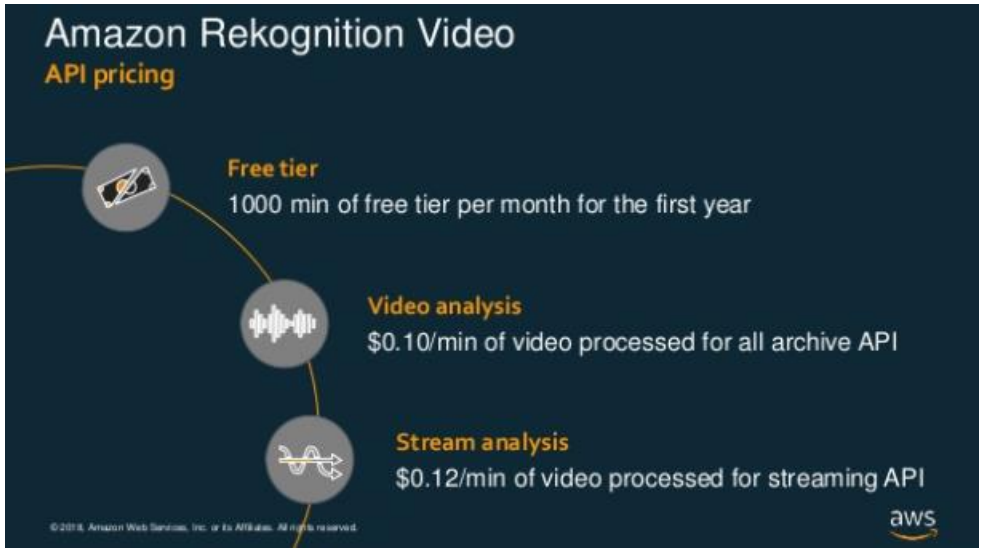
HTTPS://DOCS.AWS.AMAZON.COM/REKOGNITION/LATEST/DG/PERSONS.HTML

PUBLIC „SAFETY“



HTTPS://DE.SLIDESHARE.NET/AMAZONWEBSERVICES/HOW-TO-GET-THE-MOST-OUT-OF-AMAZON-REKOGNITION-VIDEO-A-DEEP-LEARNING-BASED-VIDEO-ANALYSIS-SERVICE-AWS-ONLINE-TECH-TALK-TITLE

WOS KOSTS?



HTTPS://DE.SLIDESHARE.NET/AMAZONWEBSERVICES/HOW-TO-GET-THE-MOST-OUT-OF-AMAZON-REKOGNITION-VIDEO-A-DEEP-LEARNING-BASED-VIDEO-ANALYSIS-SERVICE-AWS-ONLINE-TECH-TALK-TITLE

ACLU

ACT / DONATE

AMAZON: GET OUT OF THE SURVEILLANCE BUSINESS

SUPPORT THIS PETITION

59,731 Signed — 75,000 Needed

To Amazon:

Get out of the business of providing Rekognition face recognition to the government. You're powering dangerous surveillance that threatens your customers and communities. Stop padding your bottom line by selling out our civil rights.

HTTPS://ACTION.ACLU.ORG/PETITION/AMAZON-STOP-SELLING-SURVEILLANCE

We call on you to:

- Stop selling facial recognition services to law enforcement
- Stop providing infrastructure to Palantir and any other Amazon partners who enable ICE.
- Implement strong transparency and accountability measures, that include enumerating which law enforcement agencies and companies supporting law enforcement agencies are using Amazon services, and how.

Our company should not be in the surveillance business; we should not be in the policing business; we should not be in the business of supporting those who monitor and oppress marginalized populations.

Sincerely,

Amazonians

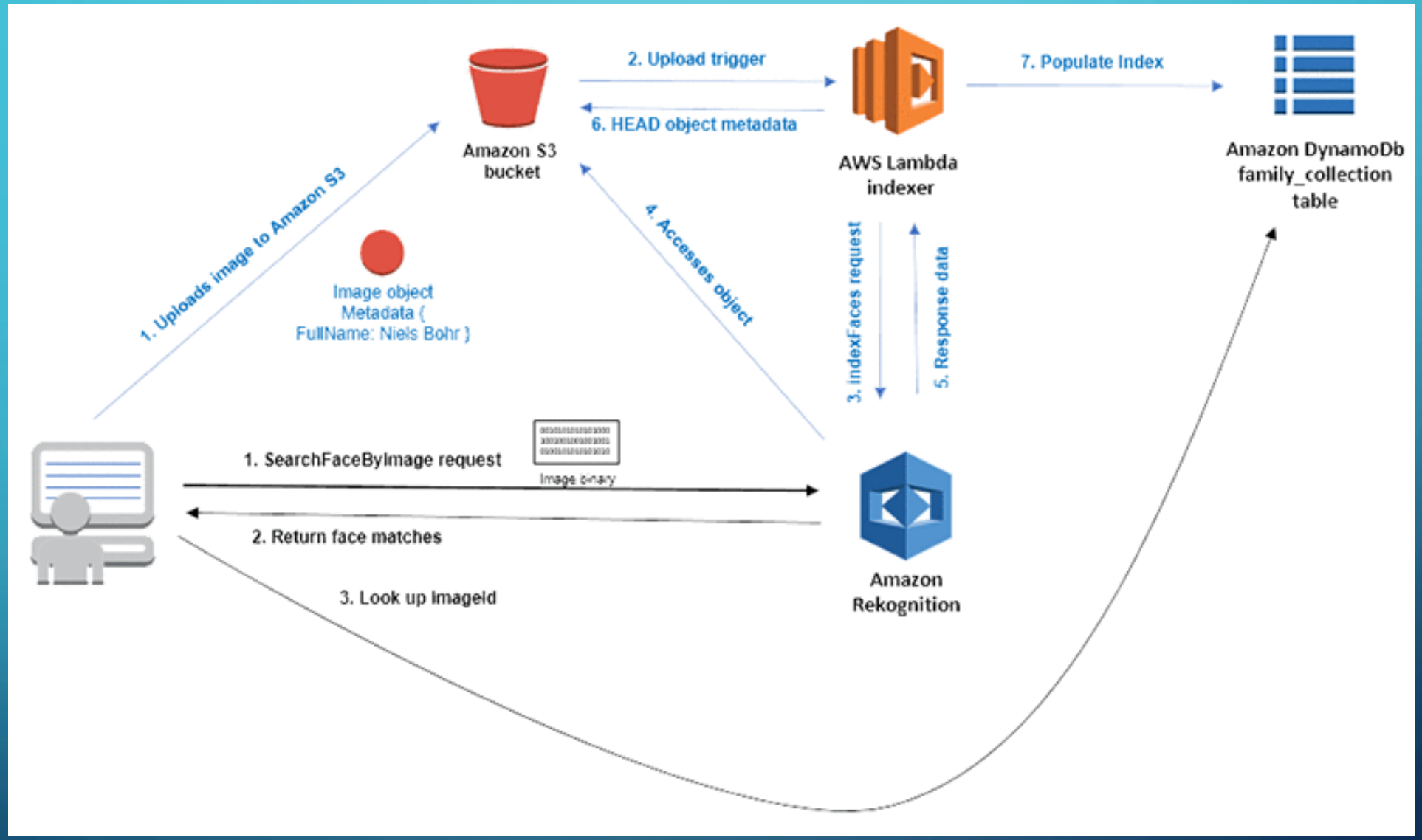
HTTPS://GIZMODO.COM/AMAZON-WORKERS-DEMAND-JEFF-BEZOS-CANCEL-FACE-RECOGNITIO-1827037509

Orlando Pulls the Plug on Its Amazon Facial Recognition Program

Amazon says its facial recognition technology can track people in a video even when their faces are not visible. Orlando, Fla., where the Police Department was using the system in a pilot program, let its contract with Amazon expire. Amazon

HTTPS://WWW.NYTIMES.COM/2018/06/25/BUSINESS/ORLANDO-AMAZON-FACIAL-RECOGNITION.HTML

WIE FUNKTIONIERT AMAZON REKOGNITION?



HTTPS://AWS.AMAZON.COM/DE/BLOGS/MACHINE-LEARNING/BUILD-YOUR-OWN-FACE-RECOGNITION-SERVICE-USING-AMAZON-REKOGNITION/

HINTER DEN KULISSEN

Convolutional Neural Net

Objectdetection

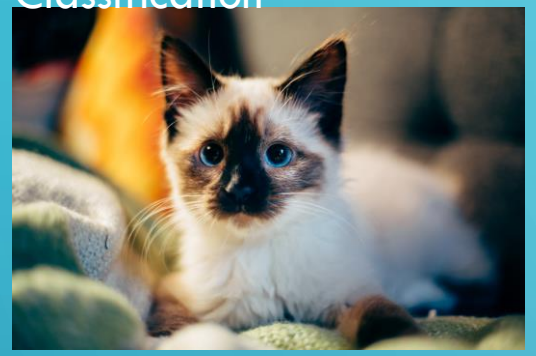
Objectidentification

CNNS

COMPUTER VISION PROBLEME

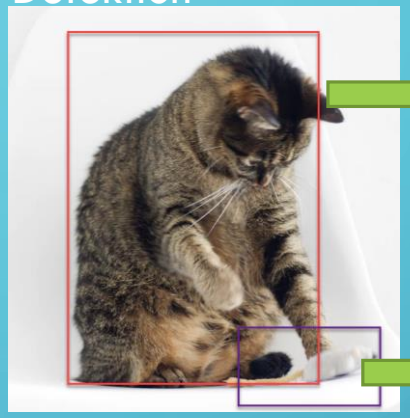
COMPUTER VISION PROBLEME

Classification



Katze?

Detektion



Katze (x=0.2,y=0.1)

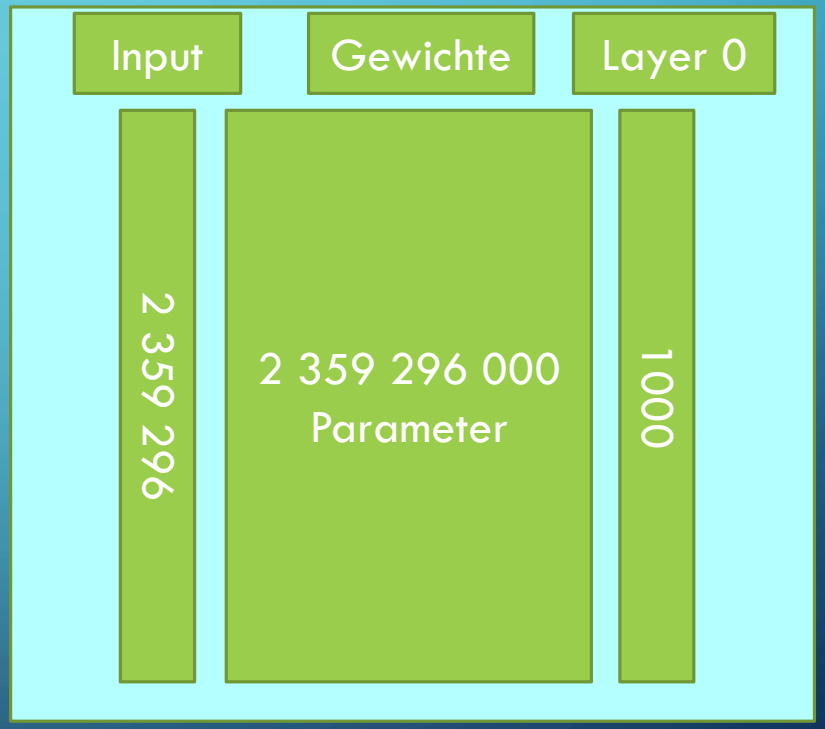
Maus (x=0.5,y=0.05)

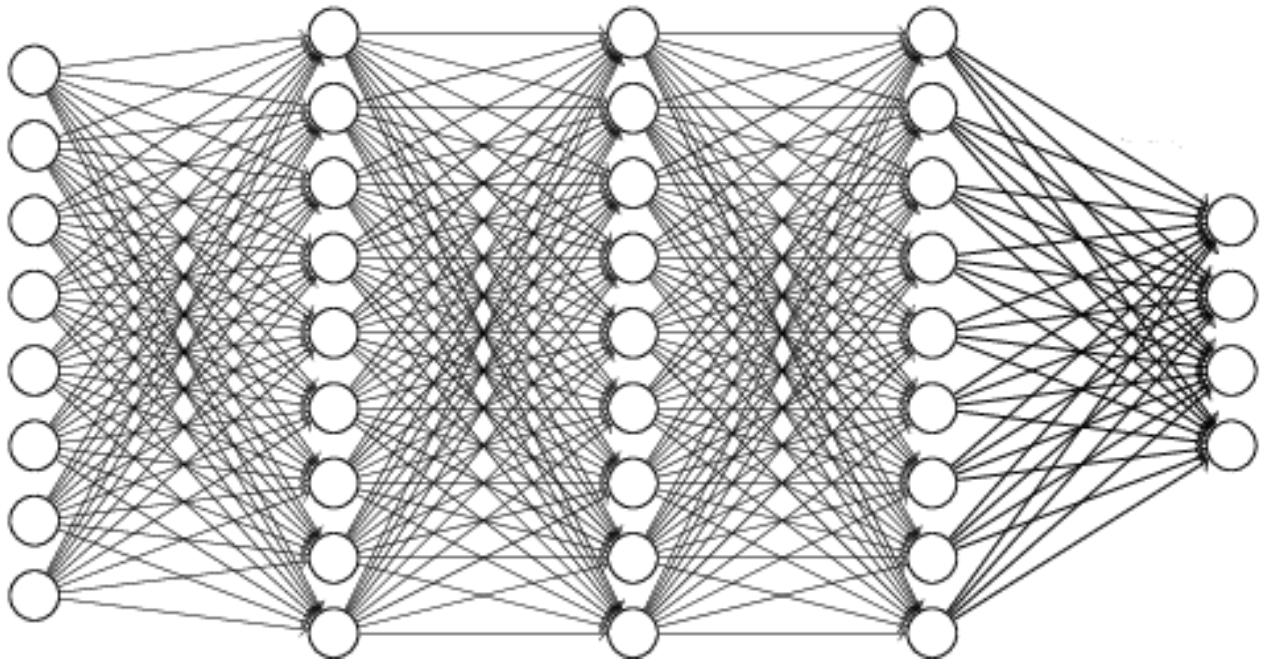
Klassifikation + Lokalisierung

Katze (x=0.25,y=0.1)

CHALLENGES

- 1024*768*3 = 2 359 296



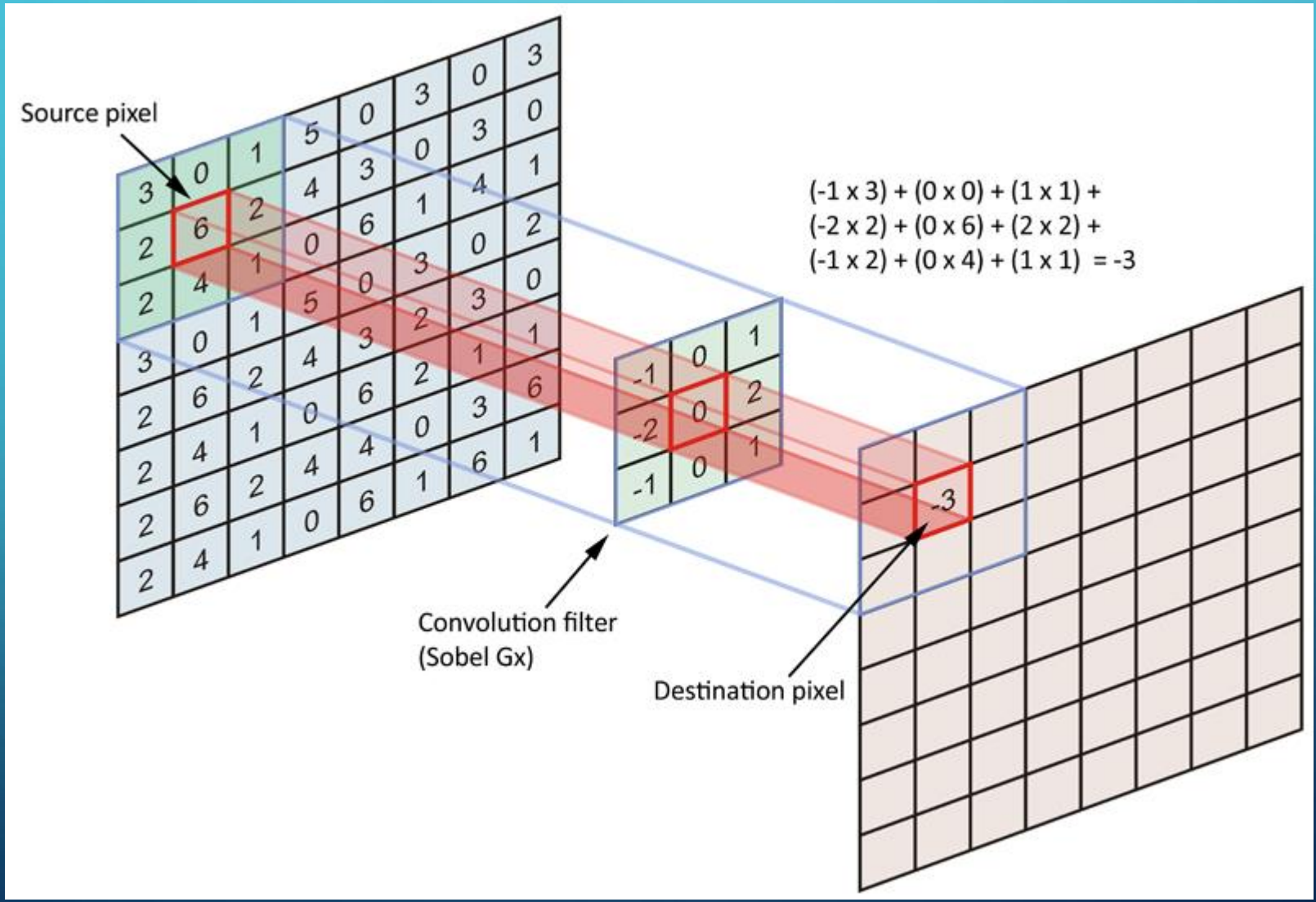


FULLY CONNECTED

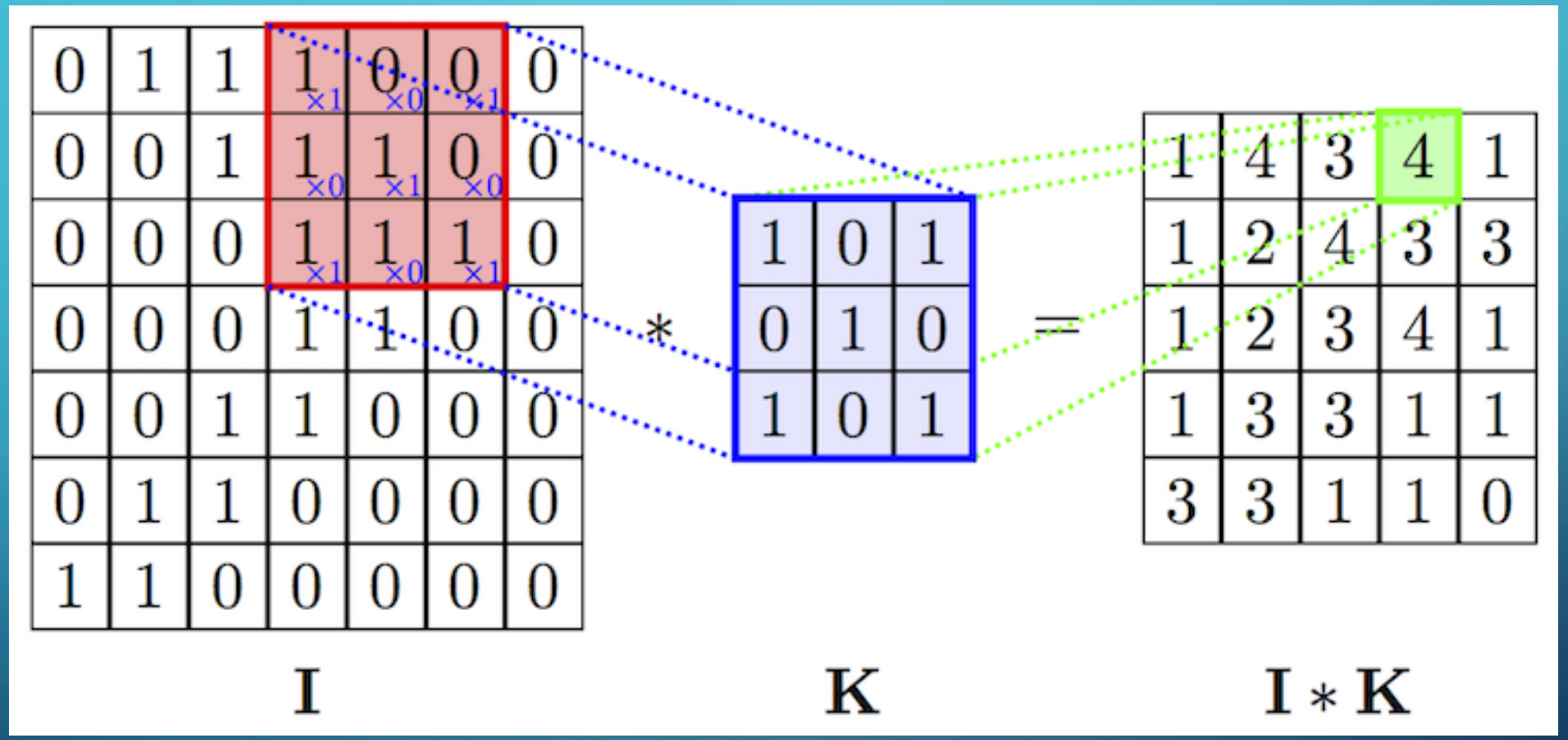
- Speicher
- Overfitting

https://math.stackexchange.com/questions/2048722/a-name-for-layered-directed-graph-as-in-a-fully-connected-neural-network

CONVOLUTION - FALTUNG



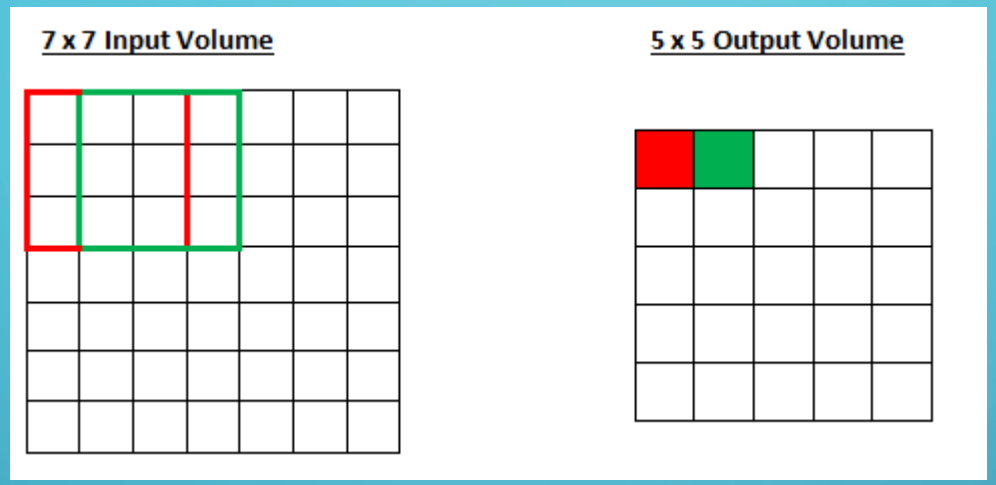
CONVOLUTION - FALTUNG

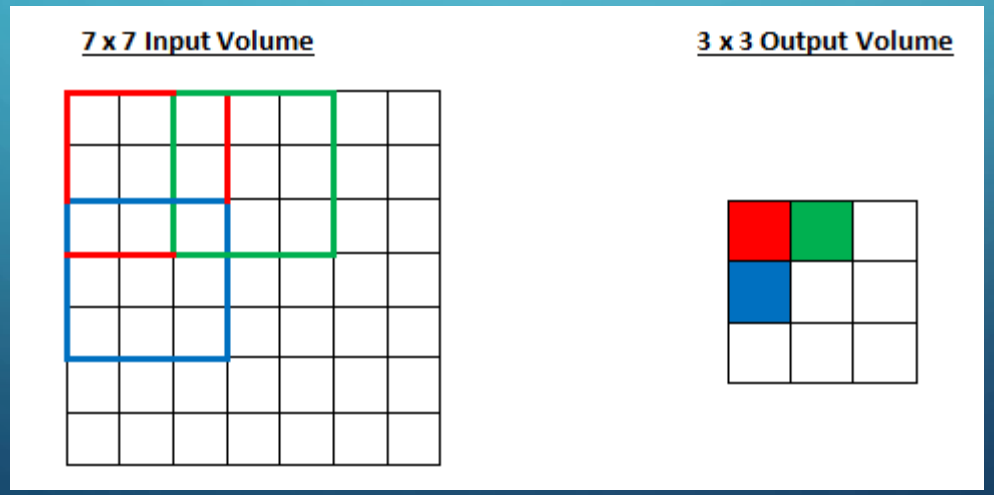


CONVOLUTION - RESULT

HTTPS://CAMBRIDGESPARK.COM/CONTENT/TUTORIALS/CONVOLUTIONAL-NEURAL-NETWORKS-WITH-KERAS/INDEX.HTML

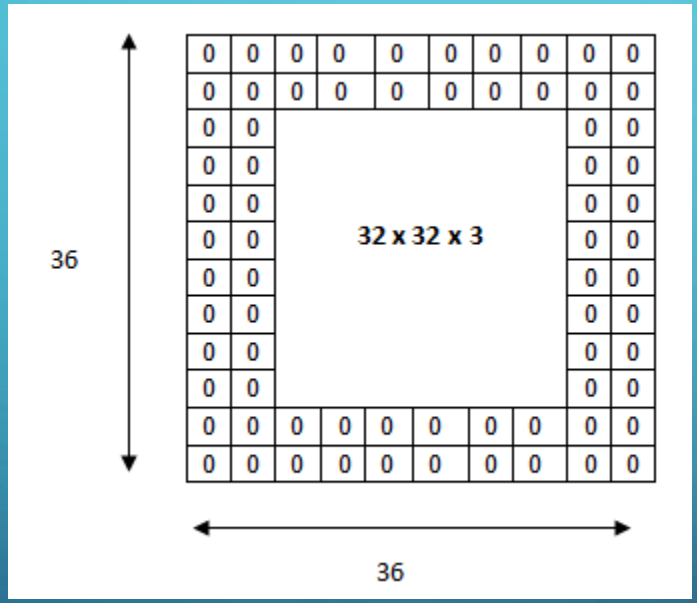
STRIDE


7 x 7 Input Volume
5 x 5 Output Volume

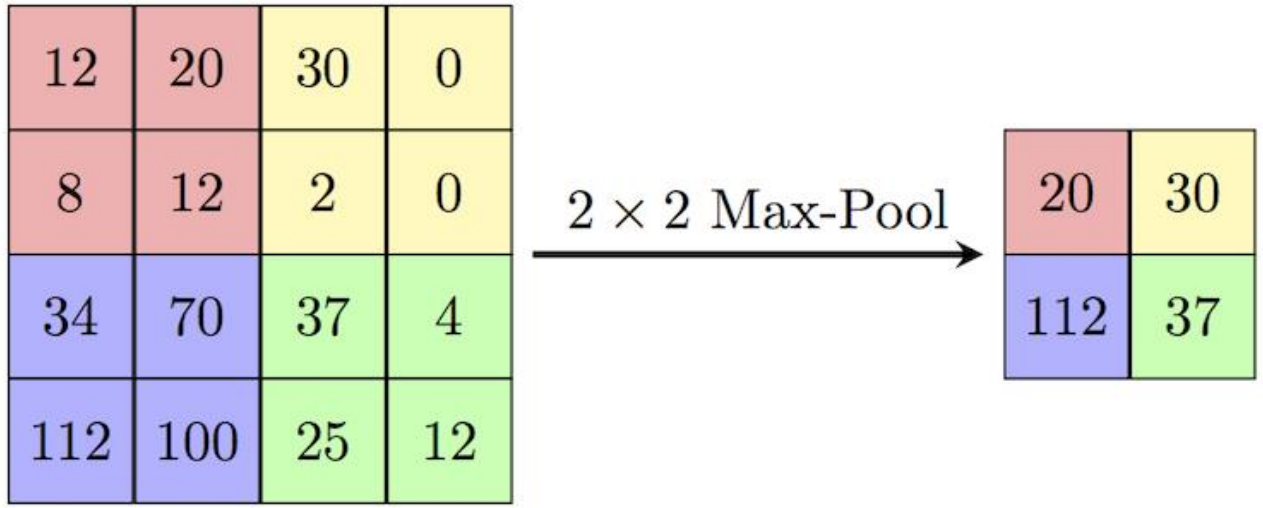

7 x 7 Input Volume
3 x 3 Output Volume

PADDING

- valid
- zero



DOWNSAMPLING - POOLING



- Unterstichproben
- Rechenlast vermindern
- weniger anfällig für Verschiebungen
- hat keine Gewichte

BEISPIEL

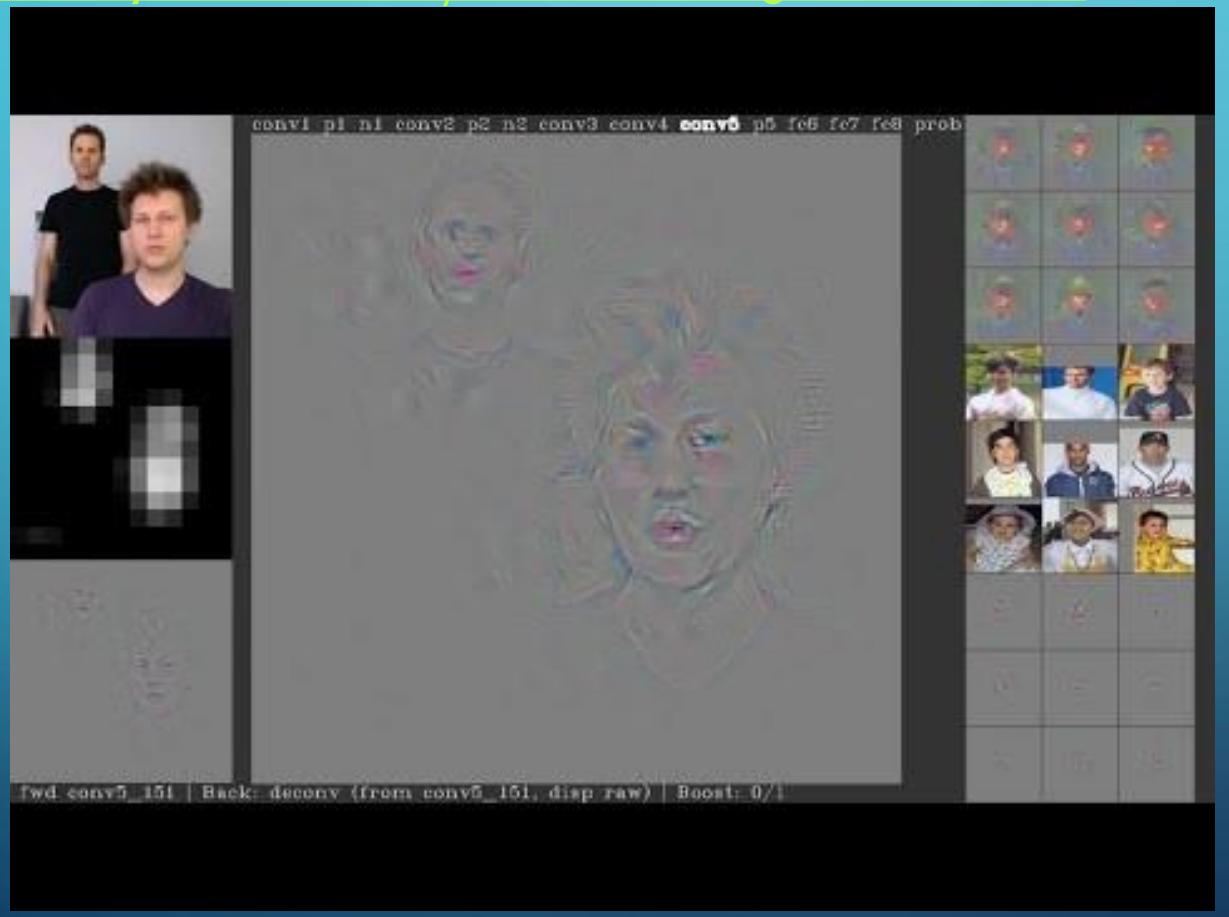
```
conv_1 = Convolution2D(conv_depth_1, (kernel_size, kernel_size), padding='same', activation='relu')(inp)
conv_2 = Convolution2D(conv_depth_1, (kernel_size, kernel_size), padding='same', activation='relu')(conv_1)
pool_1 = MaxPooling2D(pool_size=(pool_size, pool_size))(conv_2)
drop_1 = Dropout(drop_prob_1)(pool_1)

# Conv [64] -> Conv [64] -> Pool (with dropout on the pooling layer)
conv_3 = Convolution2D(conv_depth_2, (kernel_size, kernel_size), padding='same', activation='relu')(drop_1)
conv_4 = Convolution2D(conv_depth_2, (kernel_size, kernel_size), padding='same', activation='relu')(conv_3)
pool_2 = MaxPooling2D(pool_size=(pool_size, pool_size))(conv_4)drop_2 = Dropout(drop_prob_1)(pool_2)

# Now flatten to 1D, apply FC -> ReLU (with dropout) -> softmax

flat = Flatten()(drop_2)
hidden = Dense(hidden_size, activation='relu')(flat)
drop_3 = Dropout(drop_prob_2)(hidden)
out = Dense(num_classes, activation='softmax')(drop_3)
```

HTTPS://CAMBRIDGESPARK.COM/CONTENT/TUTORIALS/CONVOLUTIONAL-NEURAL-NETWORKS-WITH-KERAS/INDEX.HTML

FRAGEN?

INSIDE A NEURAL NET

- https://www.youtube.com/watch?v=AgkfIQ4IGaM



OBJECT DETECTION

Ziel:

- unterschiedliche Objekte in einem Bild klassifizieren
- eine Bounding box finden (x,y, breite, höhe)

supervised learning

FACE RECOGNITION

Input: Bild

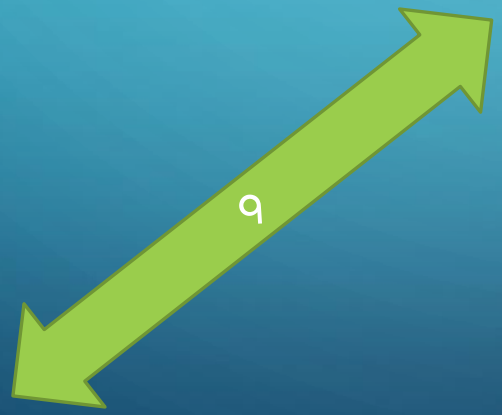
Output: Die ID, wenn es sich um ein Gesicht handelt, das in der Datenbank ist

ONE SHOT LEARNING

ONE SHOT LEARNING

- CNN Klassifikation
- d(img1, img2) = Grad des Unterschiedes von zwei Bilder
  - wenn d(img1,img2) <= t -> dieselbe Person
  - wenn d(img1,img2) >t -> unterschiedliche Personen

DIFFERENCE FUNCTION LEARNING

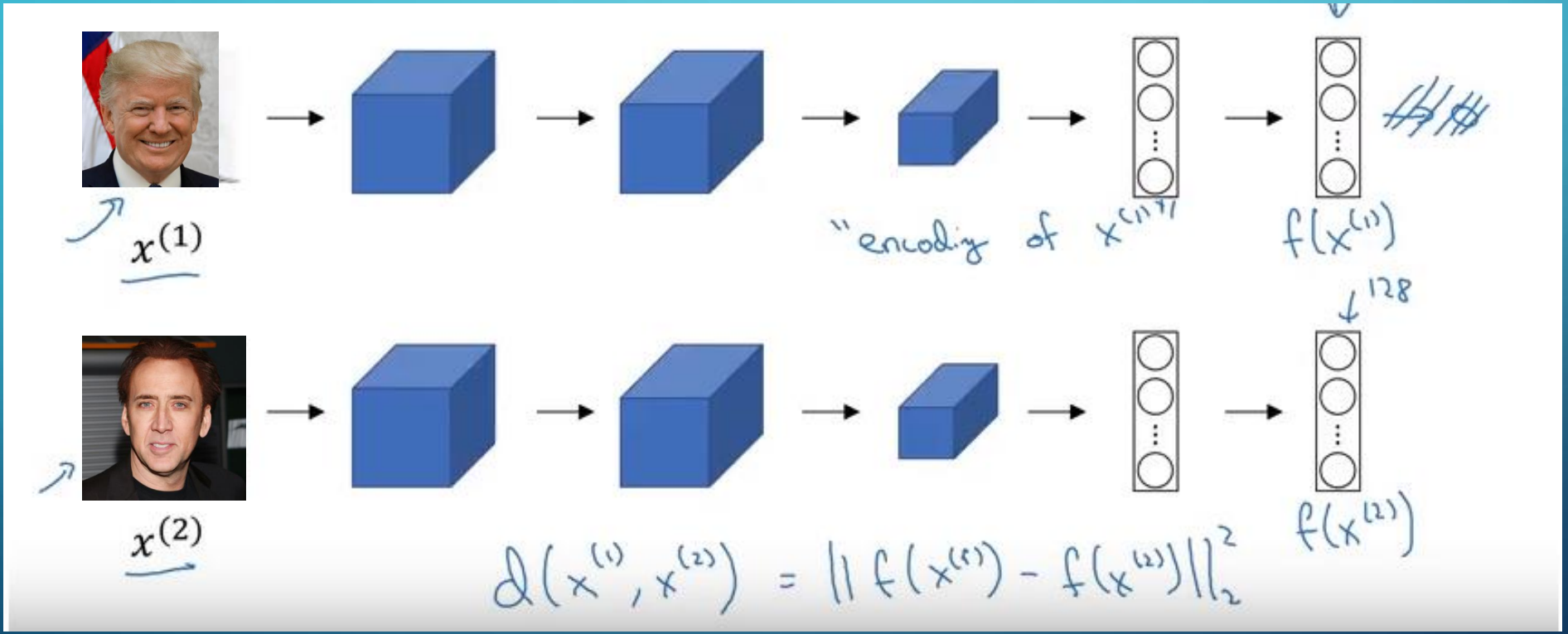


SIAMESISCHES NETZWERK

- CNN wird auf Klassifikationstask trainiert
- Outputlayer (Softmax) wird entfernt
- Netz wird zweimal instanziert
- jedes Netz erhält ein unterschiedliches Bild
- für jedes Bild wird ein Encoding (=Vektor) erstellt
- Die Vektoren werden verglichen d(img1,img2)

HTTPS://WWW.CS.TORONTO.EDU/~RANZATO/PUBLICATIONS/TAIGMAN_CVPR14.PDF

SIAMESISCHES NETZWERK



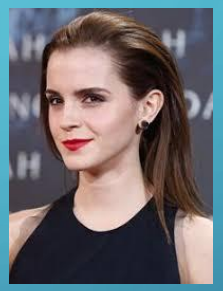
TRIPLET LOSS

- um die Encodings eines siamesischen Netzwerks unterschiedlicher Personen unterschiedlich zu machen, verwendet man eine Triplet Loss Funktion

Anchor

Positive

Anchor

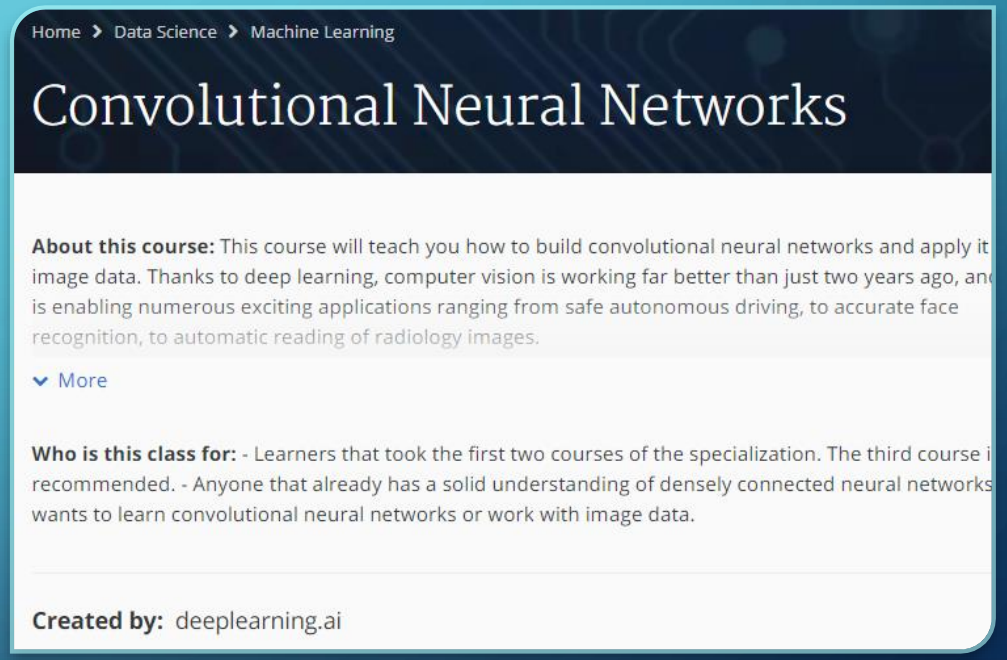


Negative

$$\mathcal{L} = max(d(a,p) - d(a,n) + margin, 0)$$

ZUM SELBER MACHEN

- https://medium.freecodecamp.org/making-your-own-face-recognition-system-29a8e728107c
- https://www.coursera.org/learn/convolutional-neural-networks



QUESTIONS?

Deep Learning Graz / Graz.ai

@hdnznger

email:
herbert.danzinger@graz.ai

PTIVE FIELDS OF SINGLE NEURONES IN THE CAT'S STRIATE CORTEX

By D. H. HUBEL* and T. N. WIESEL*

he Wilmer Institute, The Johns Hopkins Hospital and University, Baltimore, Maryland, U.S.A.

(*Received* 22 *April* 1959)

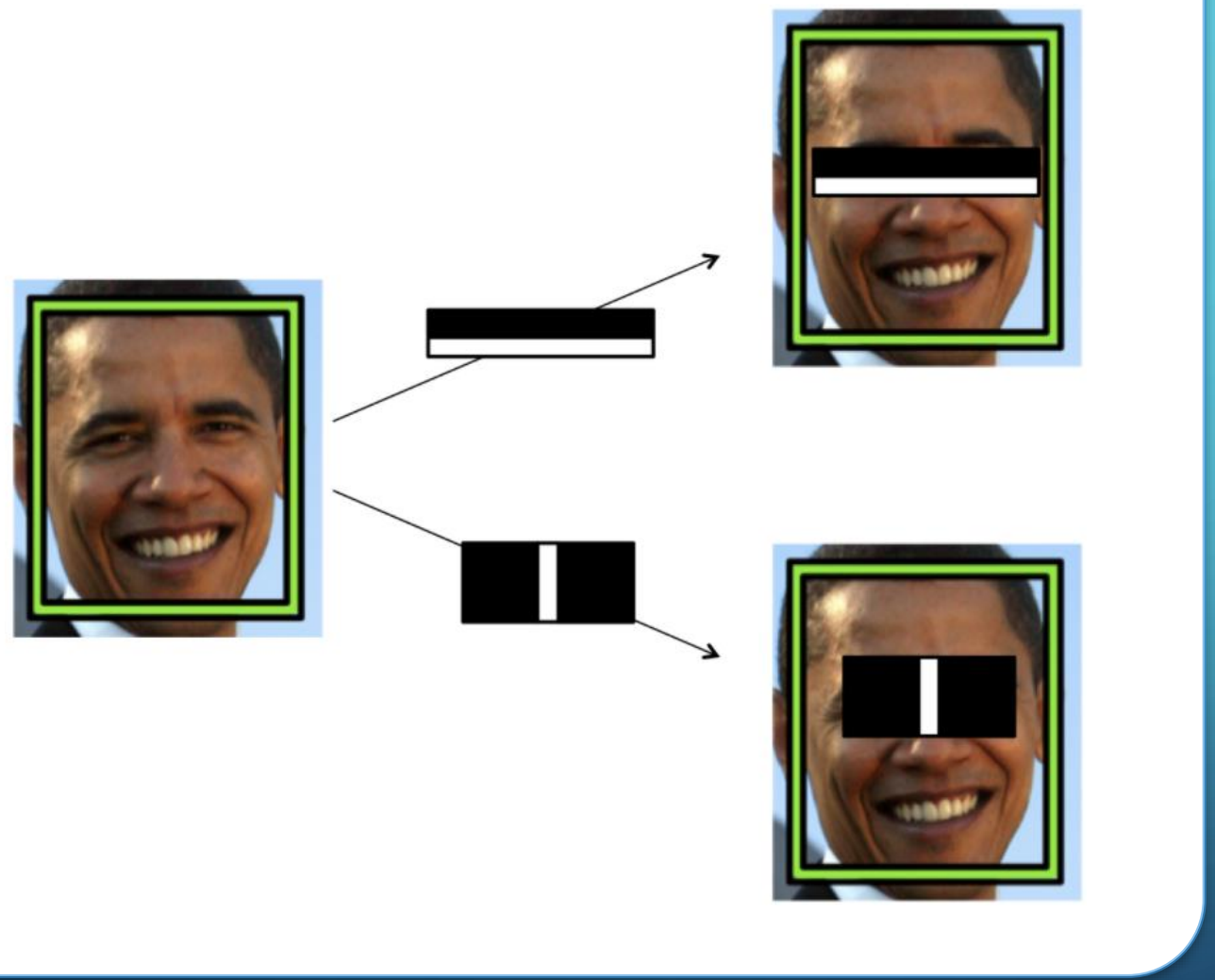
KATZENEXPERIMENTE

- manche Neuronen im visuellen Cortex reagieren auf einfache Muster
- ander wiederum auf komplexe Muster
- Abstraktionsebene/Schichtenmodell

HTTPS://WWW.NCBI.NLM.NIH.GOV/PMC/ARTICLES/PMC1363130/PDF/JPHYSIOL01298-0128.PDF

HTTPS://WWW.YOUTUBE.COM/WATCH?V=CW5PKV9RJ3O

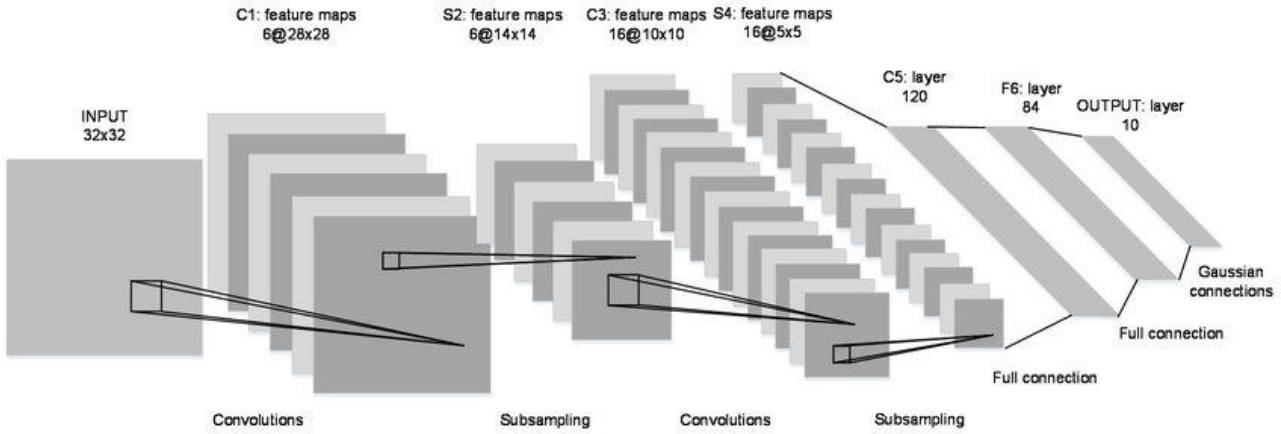
`CascadeClassifier::detectMultiScale()`



PAUL VIOLA & MICHAEL JONES

- ist auf diesem Bild ein menschliches Gesicht?
- 91,4% detection rate
- 50 false positives
- 130 Bilder, 507 Gesichter
- Nachteile:
  - Licht
  - verknittertes Foto
  - teils verdecktes Gesicht

HTTPS://DE.WIKIPEDIA.ORG/WIKI/VIOLA-JONES-METHODE



HTTPS://WWW.RESEARCHGATE.NET/FIGURE/THE-LENET-5-ARCHITECTURE-A-CONVOLUTIONAL-NEURAL-NETWORK_FIG4_321586653

LENET 5

- Yann Lecun, Leon Bottou, Yoshua Bengio, Patrick Haffner
- Erkennung handgeschriebener Ziffern auf Schecks
- Convolutional layers
- Pooling Layers
- Fully connected