VGG

UNIVERSITY OF OXFORD

Visual Geometry Group

People

Research

Publications

Demos

Data

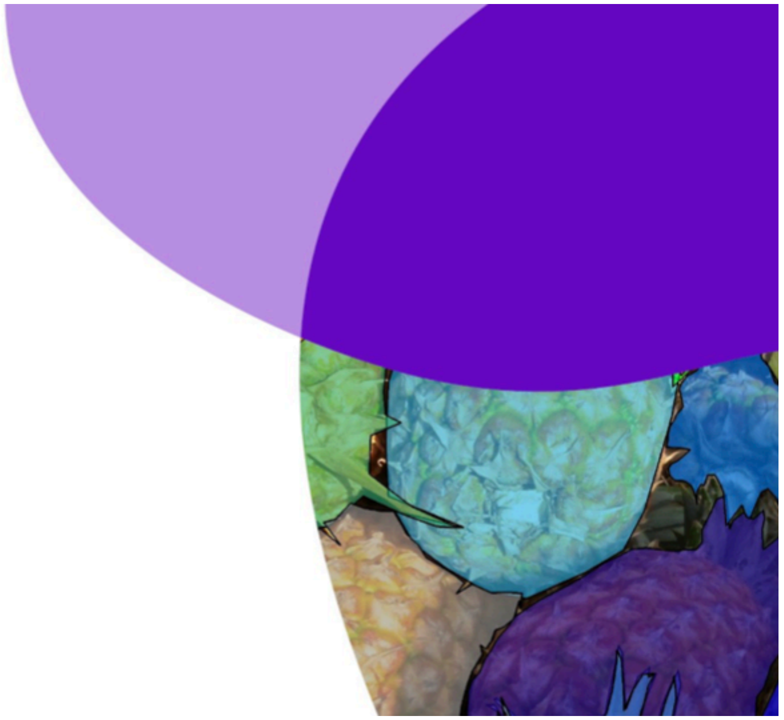
Software

Practicals

- Sudoku Associative memory
- Annealed Kullback-Leibler divergence minimization
- Set-Valued functional neural mapping
- ICA based on marginal density estimation using weighted Parzen windows
- ART, ICA, SOM, NAM, No-prop
- Introduction
- Independent component analysis note

FACEBOOK AI

Energy-Based Self-Supervised Learning



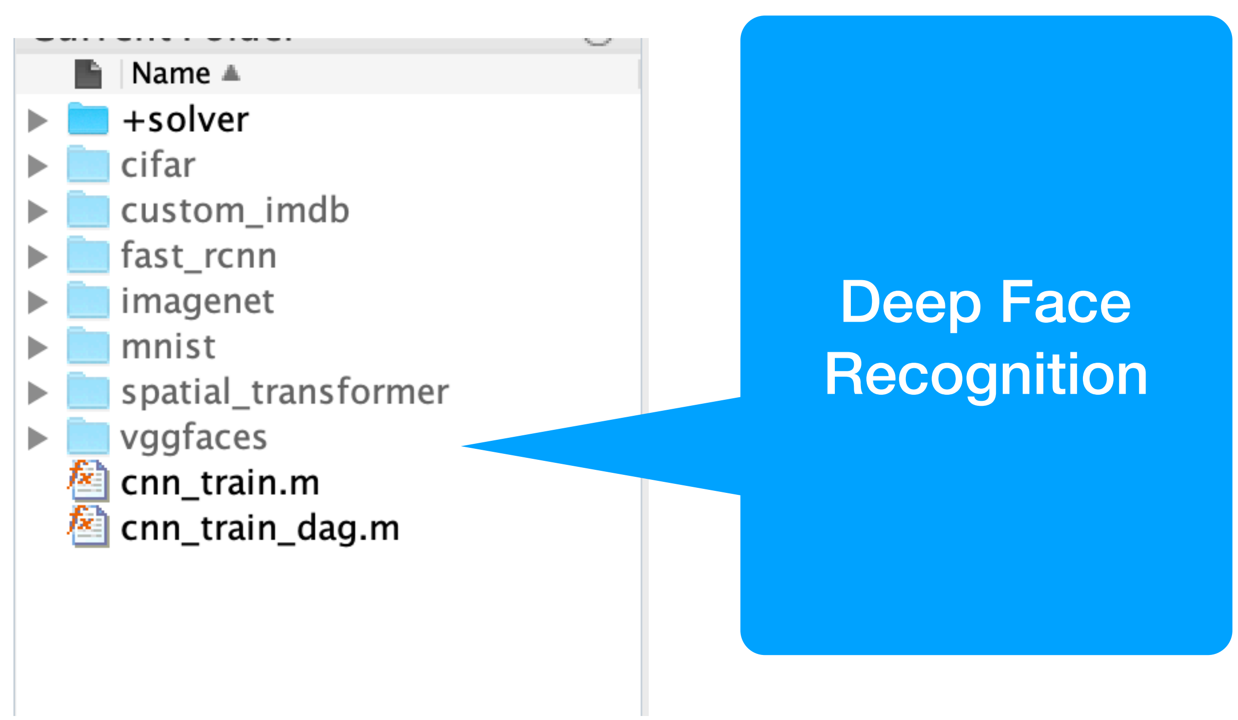
VGG-Face

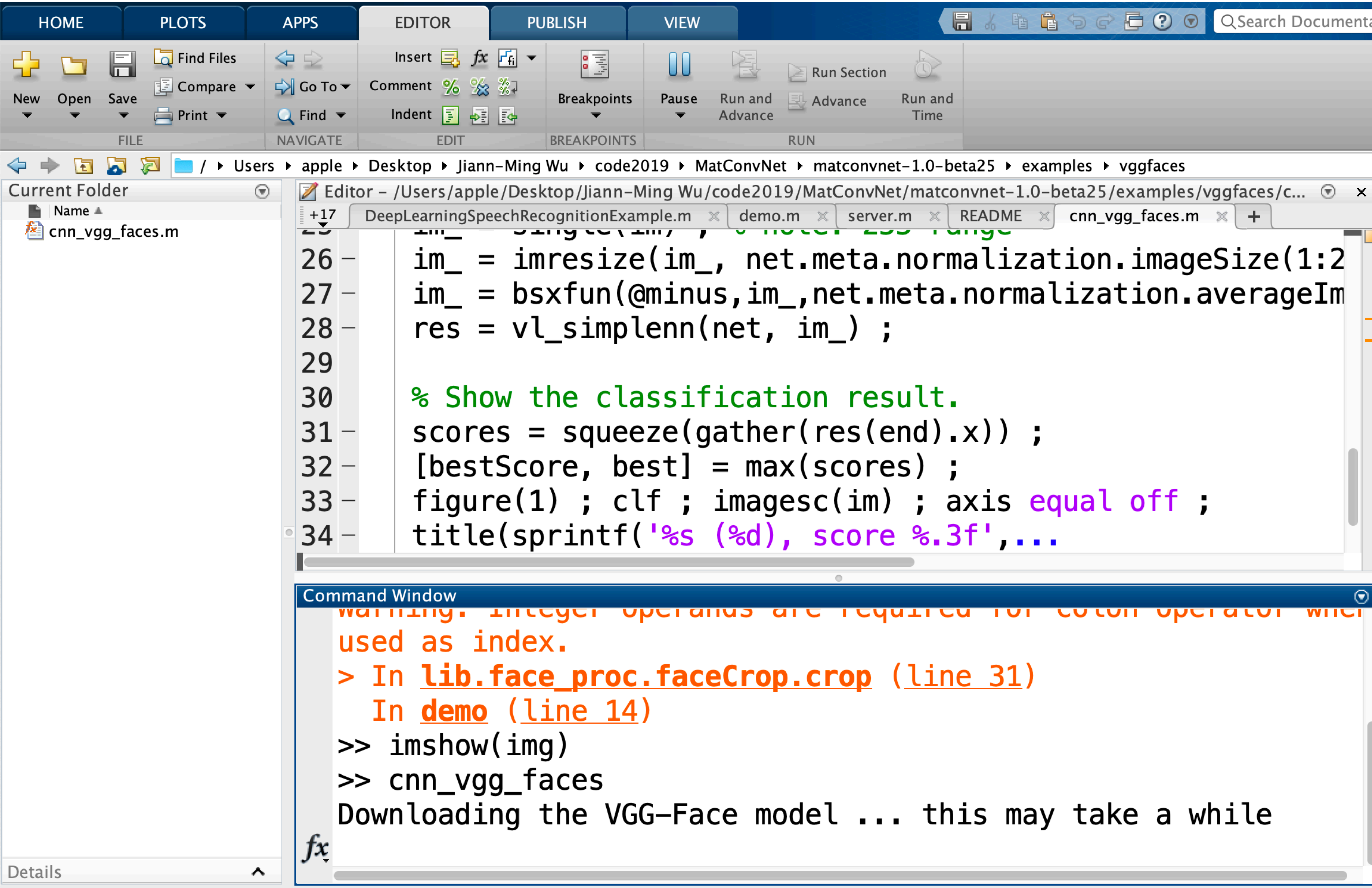
Deep Face Recognition

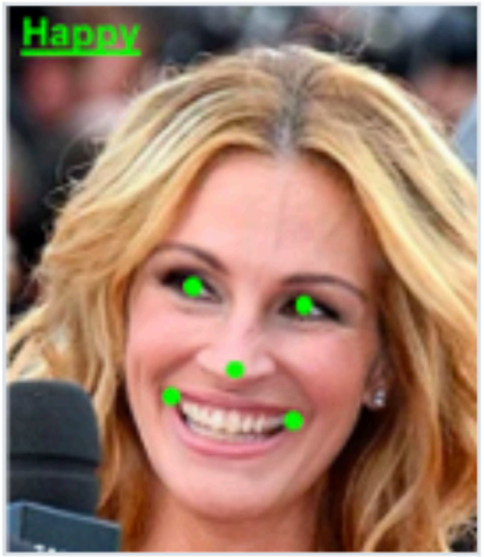
O. M. Parkhi, A. Vedaldi, A. Zisserman

British Machine Vision Conference, 2015

Download the publication :



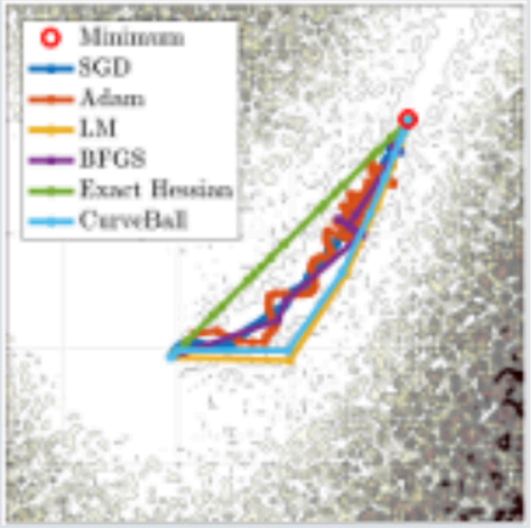

HOME
PLOTS
APPS
EDITOR
PUBLISH
VIEW
Search Documentat
New
Open
Save
Find Files
Compare
Print
FILE
Go To
Find
NAVIGATE
Insert
Comment
Indent
EDIT
Breakpoints
BREAKPOINTS
Pause
Run and Advance
Run Section
Advance
Run and Time
RUN
/ Users apple Desktop Jiann-Ming Wu code2019 MatConvNet matconvnet-1.0-beta25 examples vggfaces
Current Folder
Name
cnn_vgg_faces.m
Details
Editor – /Users/apple/Desktop/Jiann-Ming Wu/code2019/MatConvNet/matconvnet-1.0-beta25/examples/vggfaces/c...
+17
DeepLearningSpeechRecognitionExample.m
demo.m
server.m
README
cnn_vgg_faces.m
26 im_ = imresize(im_, net.meta.normalization.imageSize(1:2
27 im_ = bsxfun(@minus,im_,net.meta.normalization.averageIm
28 res = vl_simplenn(net, im_) ;
29
30 % Show the classification result.
31 scores = squeeze(gather(res(end).x)) ;
32 [bestScore, best] = max(scores) ;
33 figure(1) ; clf ; imagesc(im) ; axis equal off ;
34 title(sprintf('%s (%d), score %.3f',...
Command Window
used as index.
> In lib.face_proc.faceCrop.crop (line 31)
In demo (line 14)
>> imshow(img)
>> cnn_vgg_faces
Downloading the VGG-Face model ... this may take a while
fx



Self-supervised learning of a class-specific representation for faces

Self-supervised learning of representations that can later be used in downstream tasks such as emotion prediction or landmark regression.

Understanding and training convolutional neural networks



Small Steps and Giant Leaps: Minimal Newton Solvers for Deep Learning

Curveball is a fast second-order method that can be used as a drop-in replacement for current deep learning solvers.

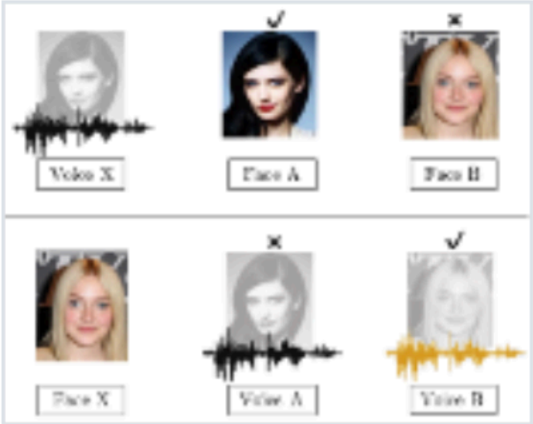
Audio-visual learning

Deep Lip Reading: A comparison of models and an online application

The goal of this work is to develop state-of-the-art models for lip reading -- visual speech recognition.

Emotion Recognition in Speech using Cross-Modal Transfer in the Wild

Transferring knowledge of emotion from faces to voices.



Seeing Voices and Hearing Faces: Cross-modal biometric matching

A network is trained to recognise faces from voices alone and vice versa.

VGG Face Descriptor

Omkar M. Parkhi, Andrea Vedaldi, Andrew Zisserman

Overview

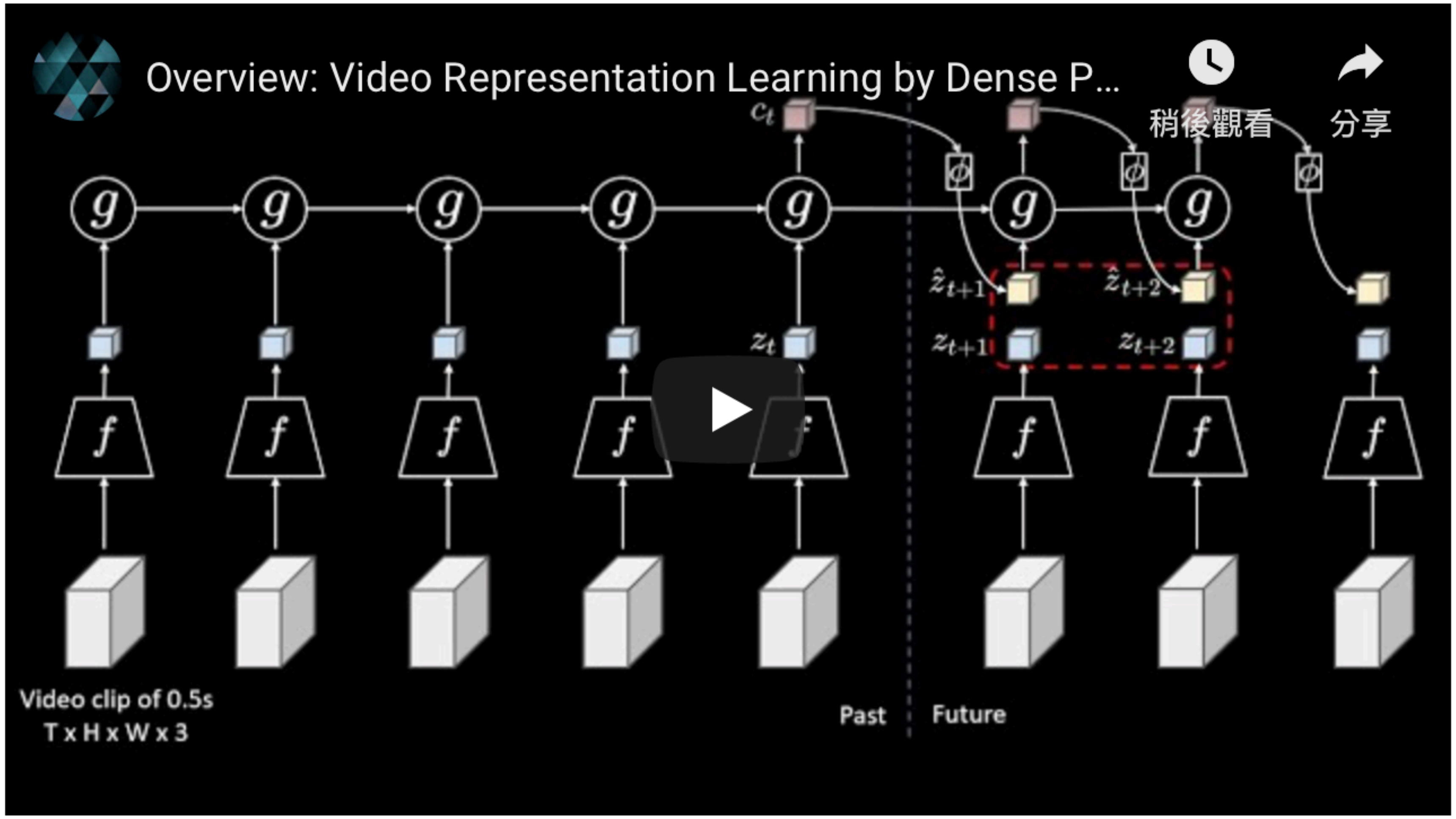
This page contains the download links for the source code for computing the **VGG-Face** CNN descriptor, described in [1].

The VGG-Face CNN descriptors are computed using our CNN implementation based on the VGG-Very-Deep-16 CNN architecture as described in [1] and are evaluated on the Labeled Faces in the Wild [2] and the YouTube Faces [3] dataset.

Additionally the code also contains our fast implementation of the DPM Face detector of [3] using the cascade DPM code of [4].

Details of how to crop the face given a detection can be found in vgg_face_matconvnet package below in class faceCrop in +lib/+face_proc directory.

The objective of this paper is self-supervised learning of spatio-temporal embeddings from video, suitable for human action recognition. We make three contributions: First, we introduce the Dense Predictive Coding (DPC) framework for self-supervised representation learning on videos. This learns a dense encoding of spatio-temporal blocks by recurrently predicting future representations; Second, we propose a curriculum training scheme to predict further into the future with progressively less temporal context. This encourages the model to only encode slowly varying spatial-temporal signals, therefore leading to semantic representations; Third, we evaluate the approach by first training the DPC model on the Kinetics-400 dataset with self-supervised learning, and then finetuning the representation on a downstream task, i.e. action recognition. With single stream (RGB only), DPC pretrained representations achieve state-of-the-art self-supervised performance on both UCF101(75.7% top1 acc) and HMDB51(35.7% top1 acc), outperforming all previous learning methods by a significant margin, and approaching the performance of a baseline pre-trained on ImageNet.


Overview: Video Representation Learning by Dense P...
稍後觀看
分享
Video clip of 0.5s
T x H x W x 3
Past
Future

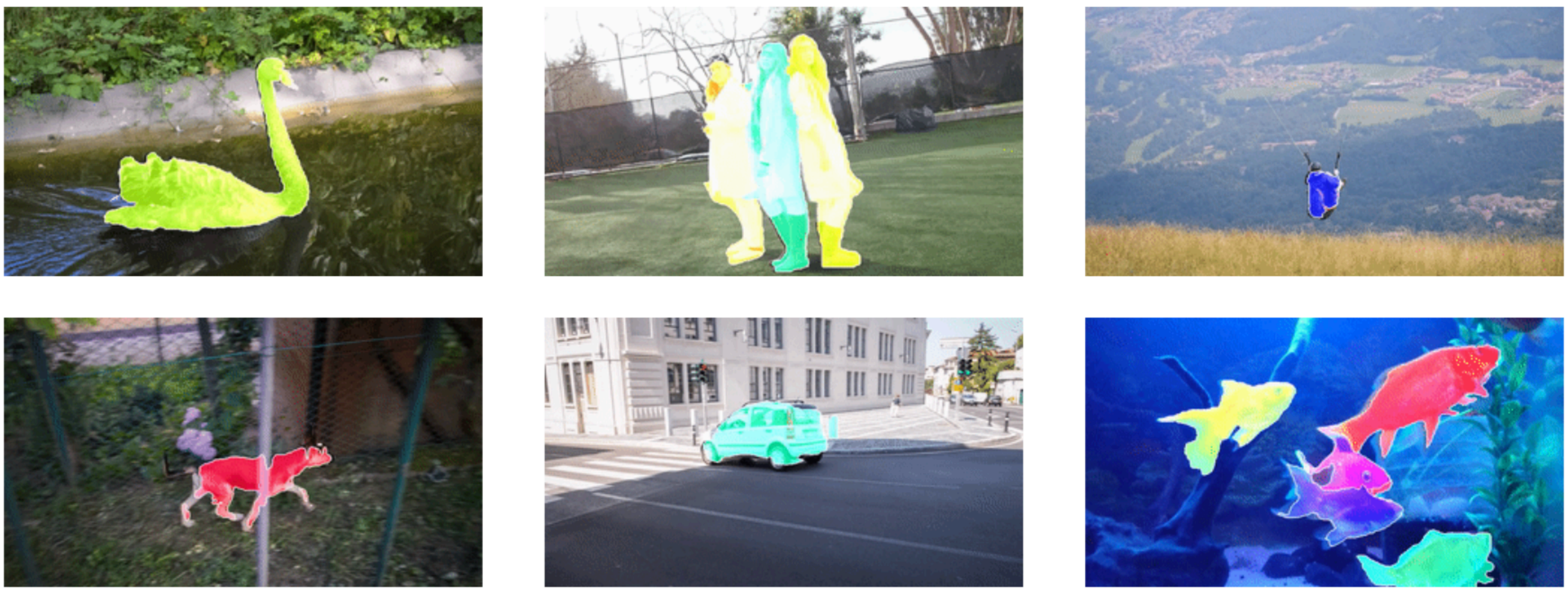
Self-supervised Learning for Video Correspondence Flow

Zihang Lai[1], Weidi Xie[2]

BMVC 2019 (Oral)

[1]Department of Computer Science, University of Oxford

[2]Visual Geometry Group, Department of Engineering Science, University of Oxford



DAVIS Semi-supervised prediction task (Given first frame)

Abstract

The objective of this paper is self-supervised learning of feature embeddings that are suitable for matching correspondences along the videos, which we term correspondence flow. By leveraging the natural spatial-temporal coherence in videos, we propose to train a "pointer" that reconstructs a target frame by copying pixels from a reference frame.

We make the following contributions: First, we introduce a simple information bottleneck that forces the model to learn robust features for correspondence matching, and prevent it from learning trivial solutions, e.g. matching based on low-level colour information. Second, to tackle the challenges from tracker drifting, due to complex object deformations, illumination changes and occlusions, we propose to train a recursive model over long temporal windows with scheduled sampling and cycle consistency. Third, we achieve state-of-the-art performance on DAVIS 2017 video segmentation and JHMDB keypoint tracking tasks, outperforming all previous self-supervised learning approaches by a significant margin. Fourth, in order to shed light on the potential of self-supervised learning on the task of video correspondence flow, we probe the upper bound by training on additional data, i.e. more diverse videos, further demonstrating significant improvements on video segmentation.

Step-by-Step Deep Learning Tutorial to Build your own Video Classification Model

Five video classification methods implemented in Keras and TensorFlow

Exploring the UCF101 video action dataset

Speech Command Recognition Using Deep Learning

R2019b

This example shows how to train a simple deep learning model that detects the presence of speech commands in audio. The example uses the Speech Commands Dataset [1] to train a convolutional neural network to recognize a given set of commands.

To run the example, you must first download the data set. If you do not want to download the data set or train the network, then you can load a pretrained network by opening this example in MATLAB® and typing `load('commandNet.mat')` at the command line. After loading the network, go directly to the last section of this example, *Detect Commands Using Streaming Audio from Microphone*.

This example uses:
Audio Toolbox
Deep Learning Toolbox

View MATLAB Command

Load Speech Commands Data Set

Download the data set from https://storage.googleapis.com/download.tensorflow.org/data/speech_commands_v0.01.tar.gz and extract the downloaded file. Set `datafolder` to the location of the data. Use `audioDatastore` to create a datastore that contains the file names and the corresponding labels. Use the folder names as the label source. Specify the read method to read the entire audio file. Create a copy of the datastore for later use.

```
datafolder = PathToDatabase;
ads = audioDatastore(datafolder, ...
    'IncludeSubfolders',true, ...
    'FileExtensions','.wav', ...
    'LabelSource','foldernames')
ads0 = copy(ads);
```