Face tracking

Across long-term sequences

Online tracking

Ground Truth

UID	Modified	Date	Annotator	Person norm.	Start Norm.	End Norm.	Duration Norm.	comment (if any)
		2001_06_30_19_00_1.webm	_title_		-1.00			
Yuji_11	-1	2001_06_30_19_00_1.webm	Yuji	Hatoyama	0:01:28	0:01:43	0:00:15	pictureonly
Minori_11	1	2001_06_30_19_00_1.webm	Minori	Hatoyama	0:01:28	0:02:32	0:01:04	picture&video
		2001_08_23_19_00_1.webm	_title_		-1.00			
Yuji_39	-1	2001_08_23_19_00_1.webm	Yuji	Koizumi	0:00:00	0:00:20	0:00:20	pictureonly
Minori_38	1	2001_08_23_19_00_1.webm	Minori	Koizumi	0:00:00	0:01:32	0:01:32	picture&video
Yuji_40	-1	2001_08_23_19_00_1.webm	Yuji	Koizumi	0:00:20	0:01:32	0:01:12	
		2001_09_19_19_00_1.webm	_title_		-1.00			
Minori_47		2001_09_19_19_00_1.webm	Minori	Koizumi	0:00:11	0:00:29	0:00:18	picture
Yuji_49		2001_09_19_19_00_1.webm	Yuji	Koizumi	0:00:11	0:00:29	0:00:18	pictureonly
Minori_48		2001_09_19_19_00_1.webm	Minori	Kanzaki	0:00:43	0:00:48	0:00:05	
Yuji_50		2001_09_19_19_00_1.webm	Yuji	Kanzaki	0:00:43	0:00:48	0:00:05	ruling party leader talks
Minori_49		2001_09_19_19_00_1.webm	Minori	Koizumi	0:00:55	0:01:01	0:00:06	
Yuji_51		2001_09_19_19_00_1.webm	Yuji	Koizumi	0:00:56	0:01:01	0:00:05	ruling party leader talks
Yuji_52	1	2001_09_19_19_00_1.webm	Yuji	Kanzaki	0:00:58	0:01:01	0:00:03	ruling party leader talks
Yuji_53	1	2001_09_19_19_00_1.webm	Yuji	Koizumi	0:01:03	0:01:04	0:00:01	ruling party leader talks

1865

545 detections

22 Faces

37%

Ground Truth

Right detections	216 (37%)
Wrong detections	38 (6.5%)
Missing detections	329 (56.4%)
Positive predictive value	85%
Sensitivity	39.6%

Name	Detections in ground truth	Detections in VC	Right Detection (True Positive)	Wrong Detection (False Positive)	Missing Detection (False Negative)	No Detection (True Negative)	Total segments	Prevalence	Accuracy	True Positive Rate
Kazuo SHII	24	13	8	5	16	1836	1865	1 %	99 %	33 %
Yukio HATOYAMA	86	32	27	5	59	1774	1886	5 %	97 %	31 %
Yasuo FUKUDA	51	13	10	3	41	1811	1865	3 %	98 %	20 %
Junichiro KOIZUMI	93	37	28	9	65	1763	1865	5 %	96 %	30 %
Takenori KANZAKI	19	2	2	0	17	1846	1865	1 %	99 %	11 %
Taro ASO	57	14	12	2	45	1806	1865	3 %	97 %	21 %
Mizuho FUKUSHIMA	33	8	7	1	26	1831	1865	2 %	99 %	21 %
Katsuya OKADA	48	5	5	0	43	1817	1865	3 %	98 %	10 %
Akihiro OHTA	12	3	3	0	9	1853	1865	1 %	100 %	25 %
Naoto KAN	88	20	18	2	70	1775	1865	5 %	96 %	20 %
Shinzo ABE	66	24	21	3	45	1798	1865	4 %	97 %	32 %
Takako DOI	0	0	0	0	0	1859	1865	0 %	100 %	0 %
Shizuka KAMEI	31	4	4	0	27	1834	1865	2 %	99 %	13 %
Ichiro OZAWA	66	26	22	4	44	1795	1865	4 %	97 %	33 %
Yukio EDANO	32	4	4	0	28	1833	1865	2 %	98 %	12 %
Sadakazu TANIGAKI	28	5	5	0	23	1837	1865	2 %	99 %	18 %
Yoshihiko Noda	62	25	23	0	39	1803	1865	3 %	98 %	37 %
Natsuo YAMAGUCHI	19	7	5	2	14	1844	1865	1 %	99 %	26 %
Seiji MAEHARA	34	7	7	0	27	1831	1865	2 %	99 %	21 %
Yoshimi WATANABE	11	2	2	0	9	1854	1865	1 %	100 %	18 %
Shintaro ISHIHARA	11	3	3	0	8	1854	1865	1 %	100 %	27 %
Ryutaro HASHIMOTO	7	2	0	2	7	1856	1865	0 %	100 %	0 %

Related work

Object Tracking with L2-RLS
Ziyang Xiao et al. (2012)

Facial shape tracking via spatio-temporal cascade shape regression
Yang et al (2015)

Shape Augmented Regression Method for Face Alignment
Wu and ji (2015)

Real-time facial landmark tracking by tree-based deformable part model based detector
Uricar and Franc (2015)

Facial Landmark Detection via Progressive Initialization
Xiao et al. (2015)

Global Supervised Descent Method
Xuehan Xiong et al.

Multi-View Constrained Local Models for Large Head Angle Facial Tracking
Rajamanoharan G, Cootes T (2015)

Online Kernel Slow Feature Analysis for Temporal Video Segmentation and Tracking — Stephan Liwicki et al. (2015)

Project-Out Cascaded Regression with an application to Face Alignment
Georgios Tzimiropoulos et al. (2015)

Development of Robust Multiple Face Tracking Algorithm and Novel Performance Evaluation Metrics for Different Background Video Sequences
Ranganatha S et al. (2018)

Face Flow
Snape et al. (2018)

Facial Landmark Detection via Progressive Initialization
— Xiao et al. (2017)

Face tracking benchmarking works

WIDER Face and Pedestrian Challenge 2018 Methods and Results
— Chen et al. (2019)

Object Tracking Benchmark
— Yi Wu et al. (2015)

A Comprehensive Performance Evaluation of Deformable Face Tracking “In-the-Wild”

— Grigorios G. Chrysos et al. (2017)

Deep Face Recognition: a Survey
— Iacopo Masi et al. (2018)

Deep Face Recognition: a Survey
— Wang et al. (2019)

Facial Emotion Recognition: A Survey and Real-World User Experiences in Mixed Reality
—Mehta et al. (2017)

The First Facial Landmark Tracking in-the-Wild Challenge: Benchmark and Results
— Shen et al (2018)

Object Tracking Methods

- Mean Shift Algorithm
  - Continuously adaptive mean shift
  - Camshift

- Linear and non-linear filtering
  - Kalman filters
  - Condensation algorithm
  - Principal Component Analysis

- Convolutional Neural Network

Face Tracking Methods

Face Detection

Facial landmark localisation

Face Tracking

Face Detection

1. Boosting Methods
2. SVM Classifier
3. Exemplar-based Techniques
4. Deep Convolutional Neural Network
5. Deformable Part Model

Facial Landmark Localisation

- Discriminative models.
  - cascaded regression
- Generative Model. Iteratively optimise
  - Active Shape Model (ASM)
  - Active Appearance Model (AAM)
- Convolutional Neural Network

MTCNN

Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks

Kaipeng Zhang et al. (2016)

20 fps

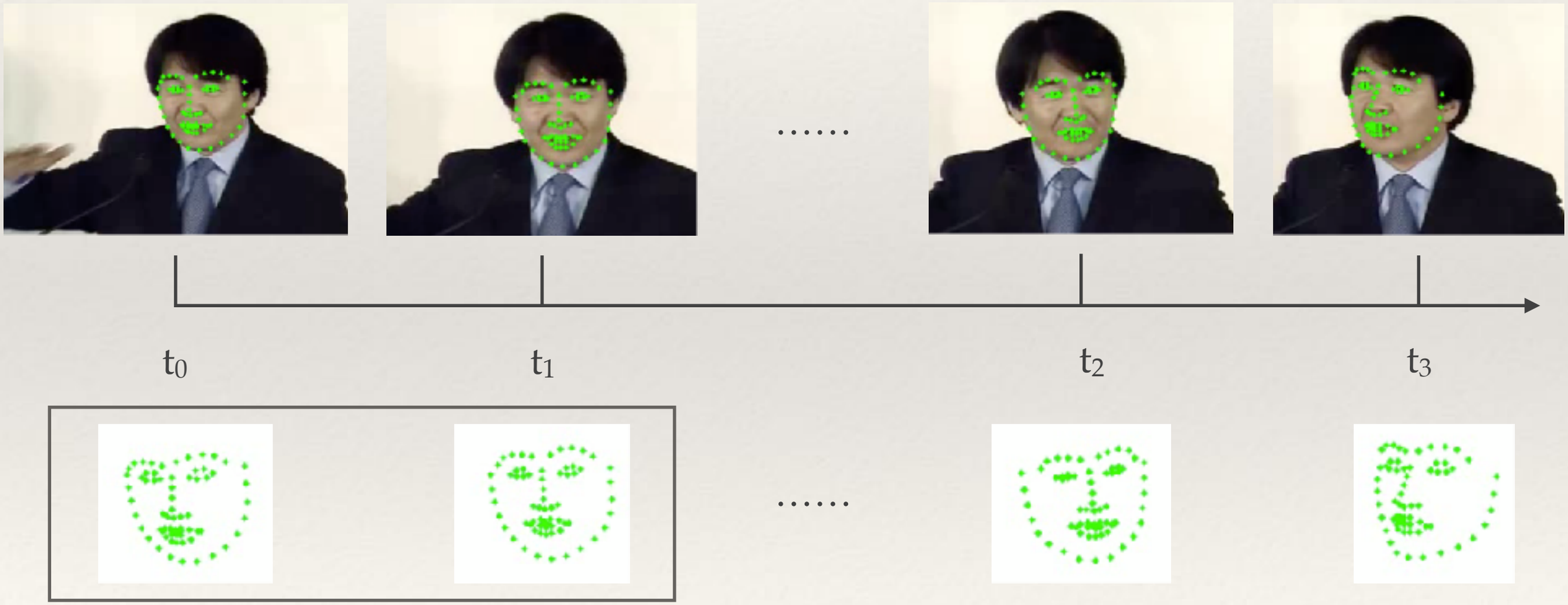
Face Tracking Methods

- Rigid 3D face tracking
- Non grid face tracking

1. **Face detection** and **facial landmark** localisation procedure at **each frame**, then using **similarity learning methods**.
2. Perform face detection in the first frame and then applies facial landmark localisation at each consecutive frame using the fitting result of the **previous frame as initialisation**, face detection re-applied in case of failure.

Face Tracking Methods

1. **Face detection** and **facial landmark** localisation procedure at **each frame**, then using **similarity learning methods**.



Similarity Learning

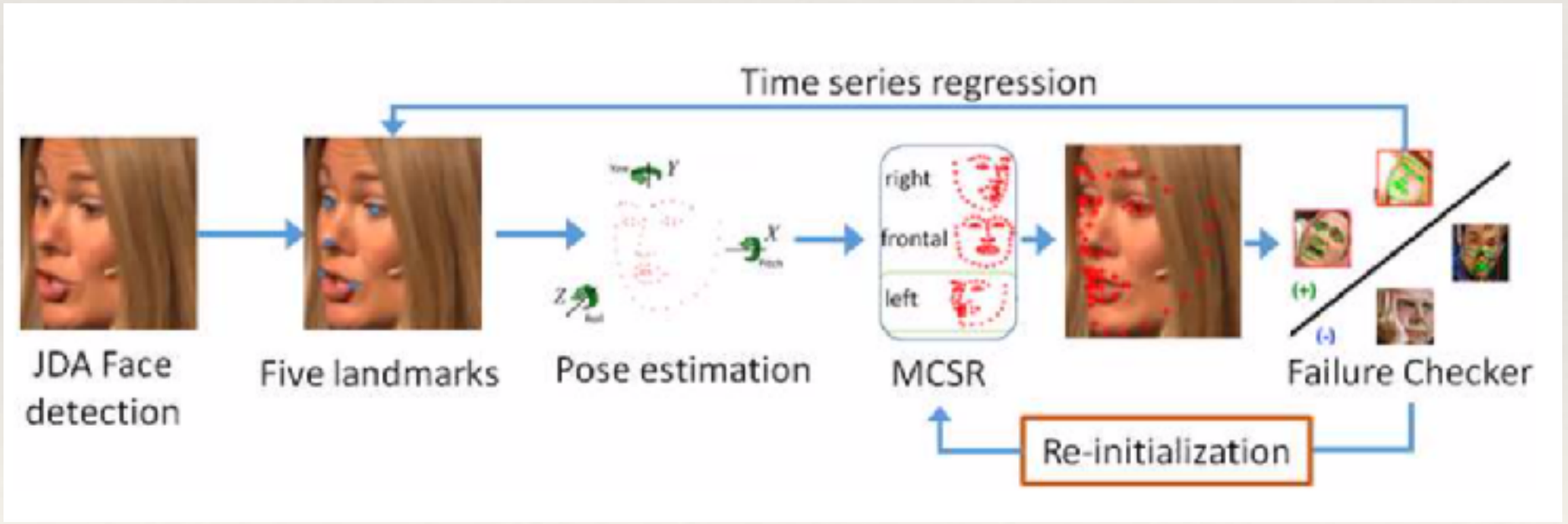
- Regression similarity learning
- Classification similarity learning
- Ranking similarity learning
- Locality sensitive hashing
- Convolutional Neural Network

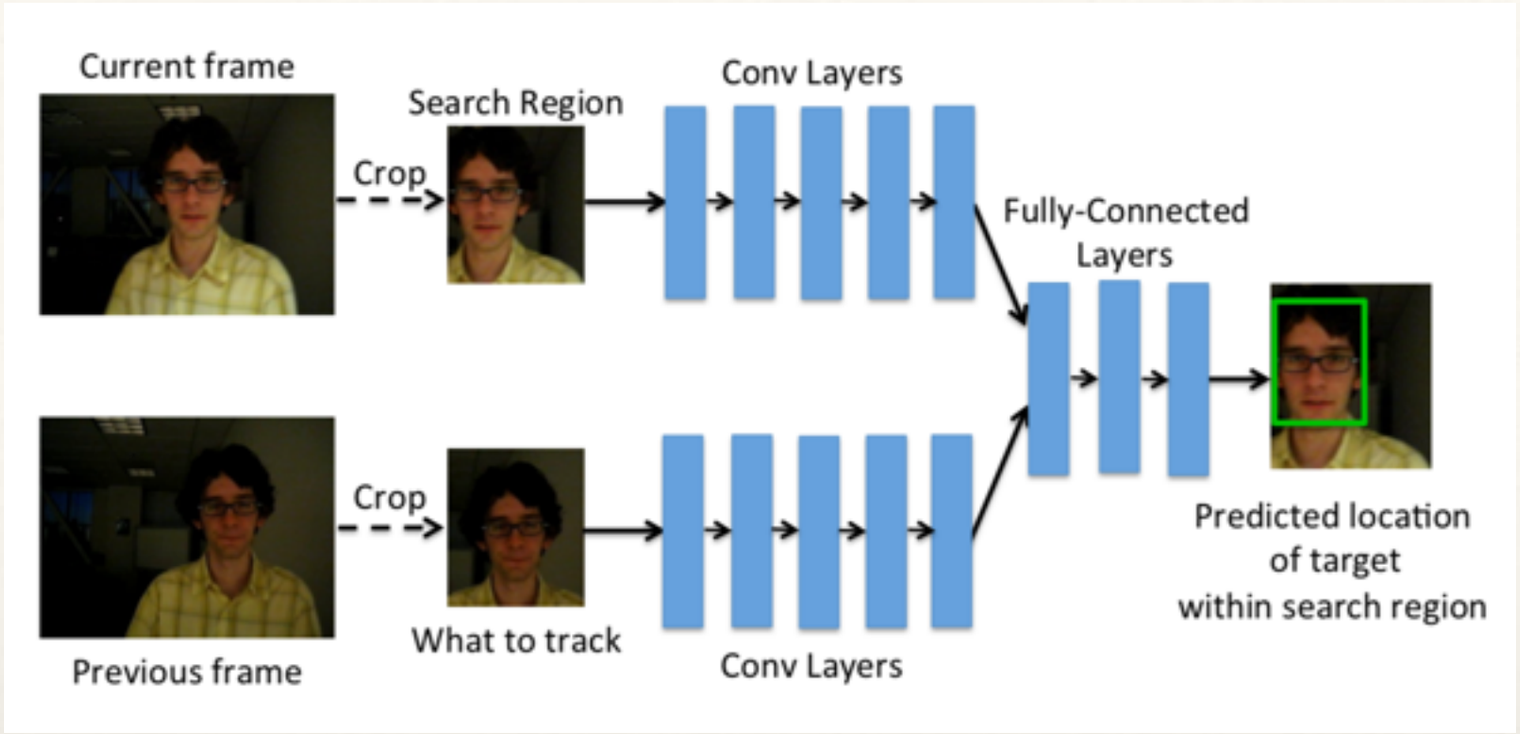
Face Tracking Methods

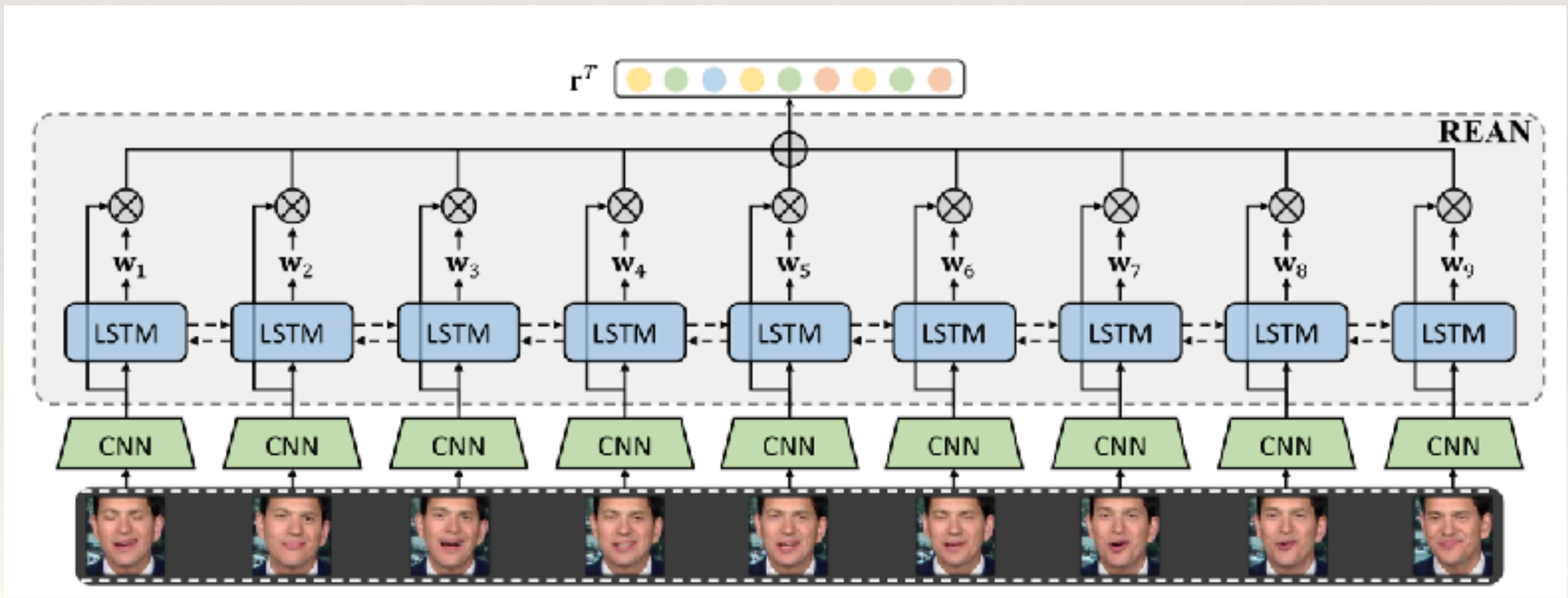
Facial shape tracking via spatio-temporal cascade shape regression
Yang et al. (2015)

300 Videos in the Wild (300-VW) Challenge & Workshop

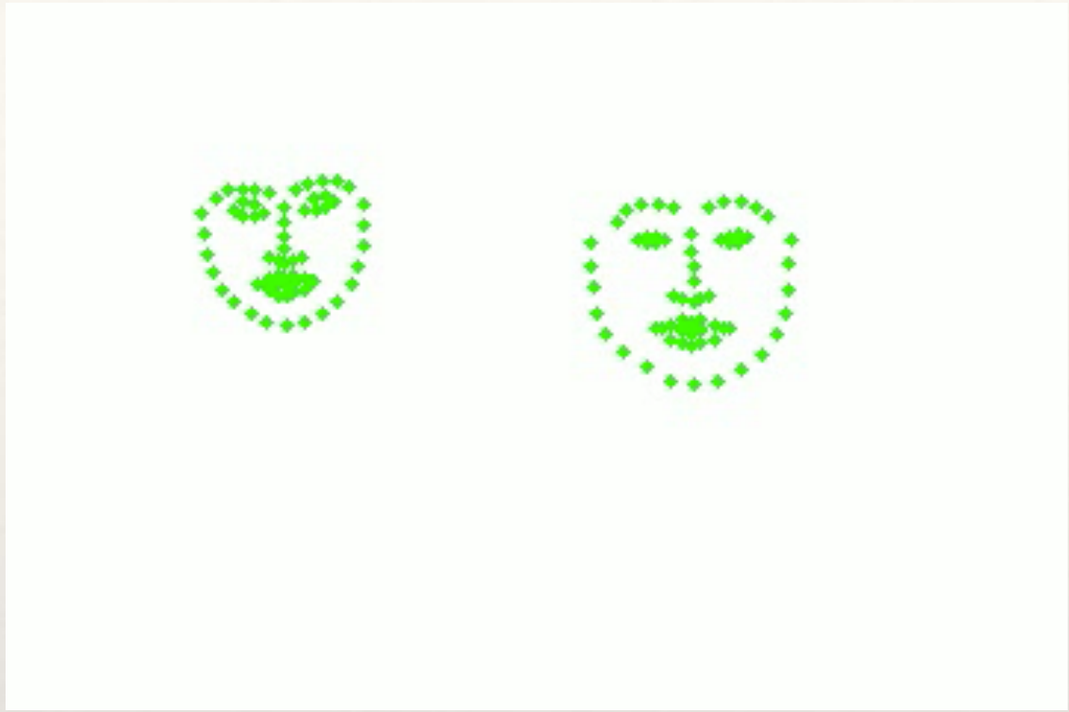
Yang et al (2015a)	0.791	2.400	0.788	0.322	0.710	4.461




Current frame
Search Region
Crop
Conv Layers
Fully-Connected Layers
Predicted location of target within search region
Crop
What to track
Previous frame
Conv Layers


$\mathbf{r}^T$
REAN
$\mathbf{w}_1$
$\mathbf{w}_2$
$\mathbf{w}_3$
$\mathbf{w}_4$
$\mathbf{w}_5$
$\mathbf{w}_6$
$\mathbf{w}_7$
$\mathbf{w}_8$
$\mathbf{w}_9$
LSTM
CNN

Face Tracking Methods

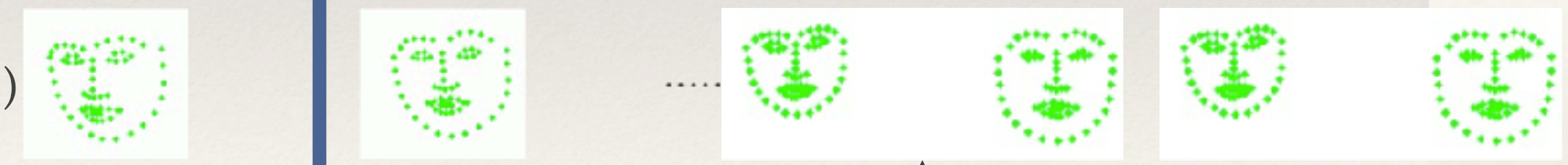


1. Face shape will not change abruptly between the consecutive frames on video
2. No illumination, occlusion problem
3. Follow what’s person doing

Face Tracking Strategy

Face Detection	Facial landmark localisation	MTCNN
Face Tracking	Regression / (LSTM)	
Face Identification / Recognition	Arcface	

(5 landmarks)

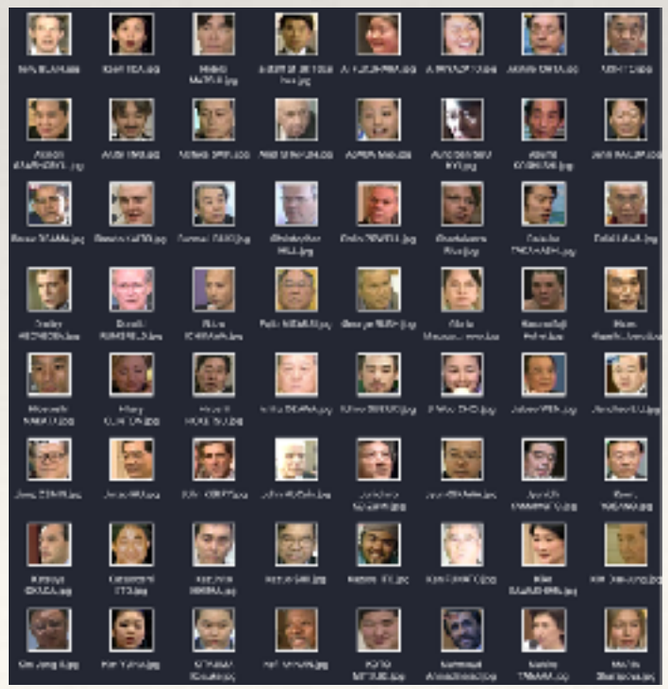


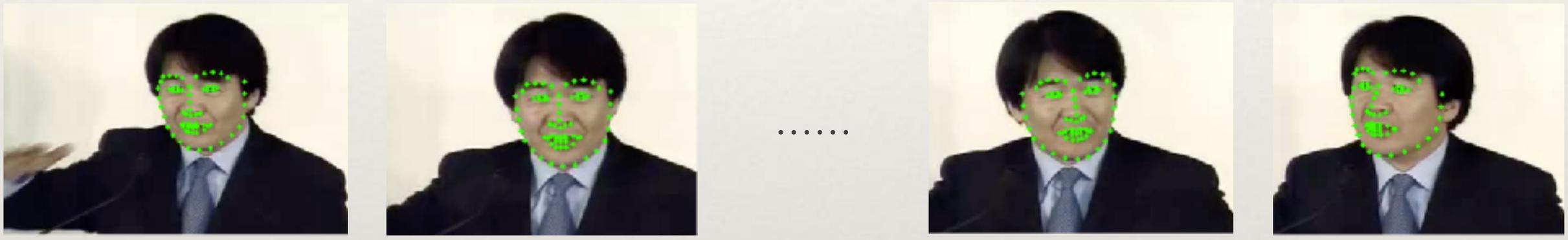
LSTM

Face Detection

Face Tracking

Face Identification




.....
.....
t_0
t_1
t_2
t_3
.....

Regression-based facial

One major limitation of recent model-based regression methods, is that they may be easily trapped by a local optimum if the starting shape is far away from the ground-truth shape.

72-points shapes

Metric Learning

Facial behaviour analysis, lip reading, surveillance, human-computer and human-robot interaction etc., require accurate continuous tracking of the facial landmarks.

the current practise regarding deformable face tracking includes the combination of a generic face detection and generic facial landmark localisation technique

face shape will not change abruptly between the consecutive frames on video

expression, illumination, occlusion, pose

wide angle changes

Verification

General Data info

Time line