DeepFakes Detection



Ruben Tolosana

ruben.tolosana@uam.es

BiDA Lab

Biometrics & Data Pattern Analytics Lab

UAM

Universidad Autónoma de Madrid

What is it?

Media Forensics has attracted a lot of attention in the last years in part due to the increasing concerns around DeepFakes.

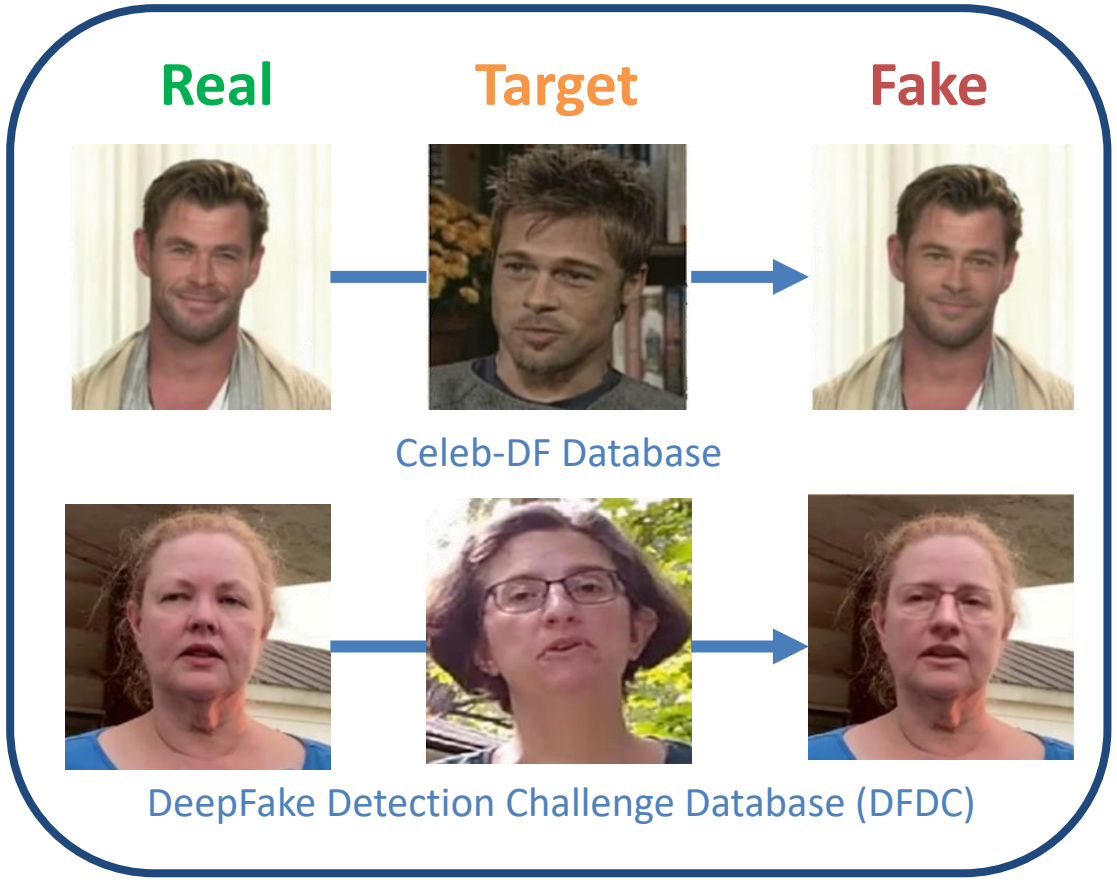
Real Video
(Robert de Niro)

DeepFake Video
(Al Pacino)

- R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales and J. Ortega-Garcia, "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection," Information Fusion, vol. 64, pp. 131-148, 2020.

What is it?

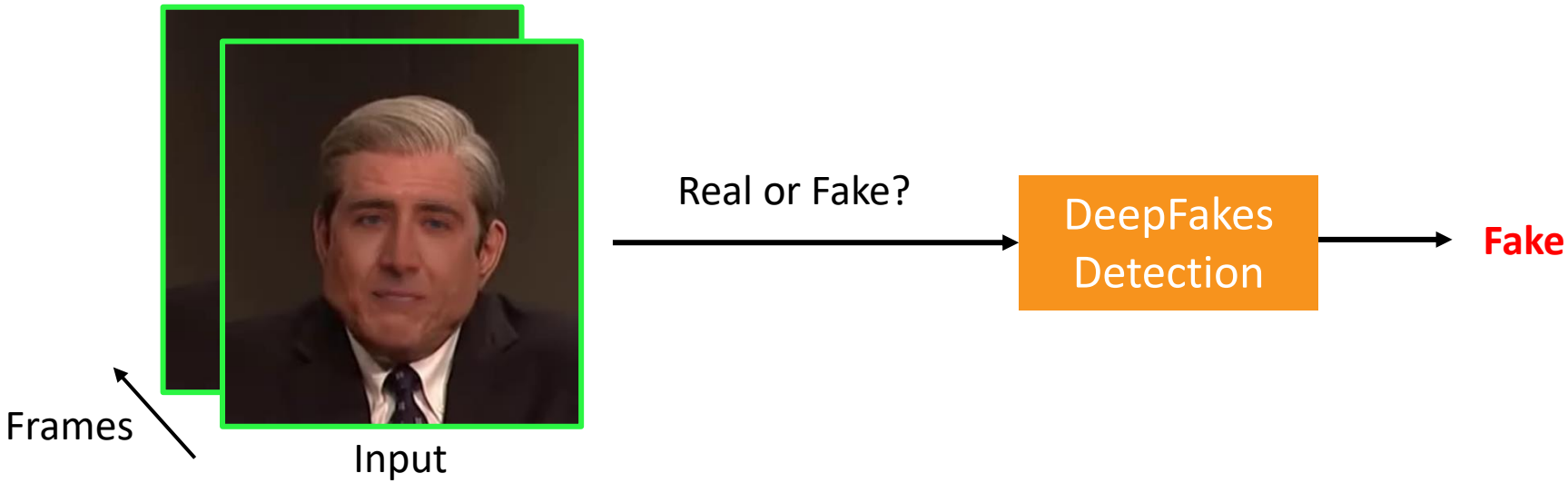
DeepFake (Identity Swap) is referred to a deep learning based technique able to create fake videos by swapping the face of a person by the face of another person.



- R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales and J. Ortega-Garcia, "DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection," Information Fusion, vol. 64, pp. 131-148, 2020.

DeepFakes Detection

Detection of digital face manipulations.



- R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales and J. Ortega-Garcia, “DeepFakes and Beyond: A Survey of Face Manipulation and Fake Detection,” Information Fusion, vol. 64, pp. 131-148, 2020.

TASK 1

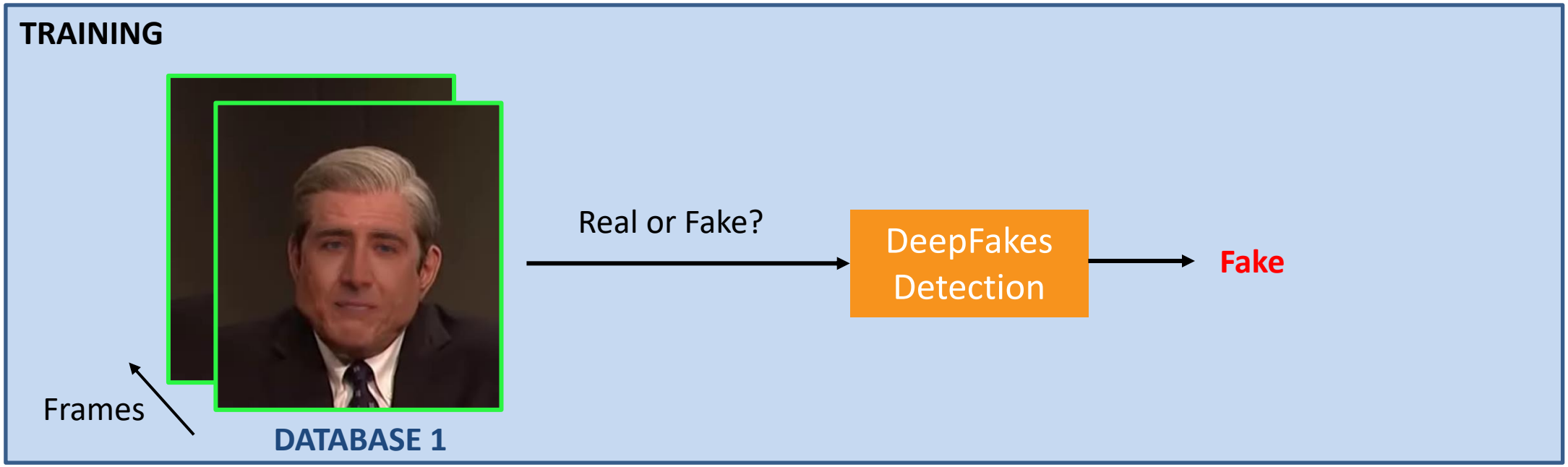
DeepFakes Detection: Intra-Database Analysis

Task 1 - DeepFakes Detection: Intra-Database Analysis

Intra-database analysis: we are going to train and evaluate the systems using the same database.

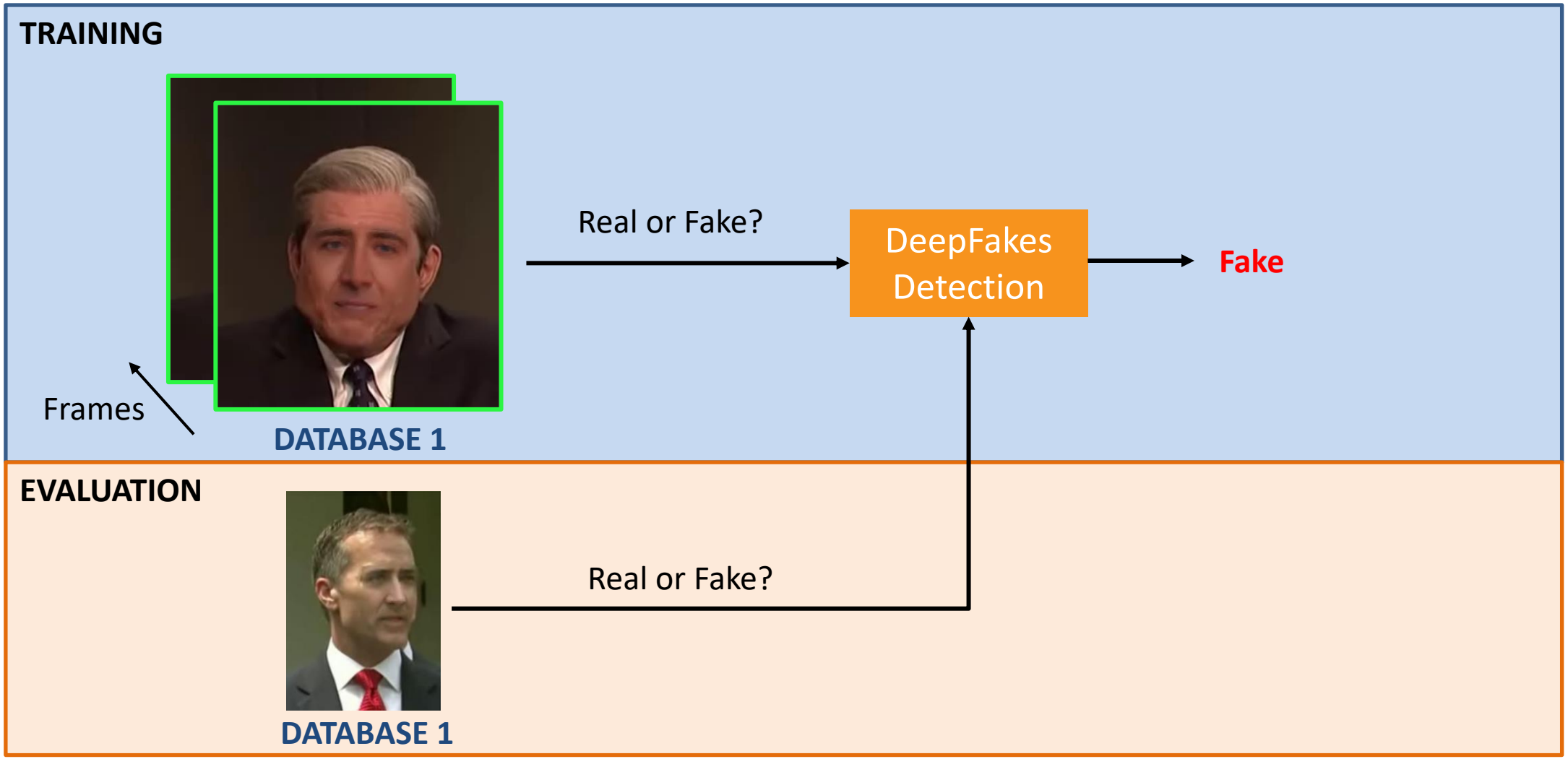
Task 1 - DeepFakes Detection: Intra-Database Analysis

Intra-database analysis: we are going to train and evaluate the systems using the same database.



Task 1 - DeepFakes Detection: Intra-Database Analysis

Intra-database analysis: we are going to train and evaluate the systems using the same database.



Task 1 - DeepFakes Detection: Intra-Database Analysis

UADFV database: 49 real and fake videos (each) using the FakeAPP software application (Computer Graphics).

- Y. Li, M. Chang, and S. Lyu, “In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking,” in Proc. IEEE Int. Workshop on Information Forensics and Security, 2018.

Task 1 - DeepFakes Detection: Intra-Database Analysis

Experimental protocol:

- Development: this dataset is considered only for the training of the systems.
  - Real folder: 38 videos.
  - Fake folder: 38 videos.

- Evaluation: this dataset is only considered for the final evaluation of the system (not for training!!!).
  - Real folder: 11 videos.
  - Fake folder: 10 videos.

For each video, we provide 10 frames.

Task 1 - DeepFakes Detection: Intra-Database Analysis

The task is to classify each image as real or fake independently, NOT VIDEOS. Therefore, for each image you will have to provide a classification score and a decision.

Input Image

Real or Fake?

DeepFakes Detection

Score: 0.93 (Real)

Input Image

Real or Fake?

DeepFakes Detection

Score: 0.27 (Fake)

TASK 2

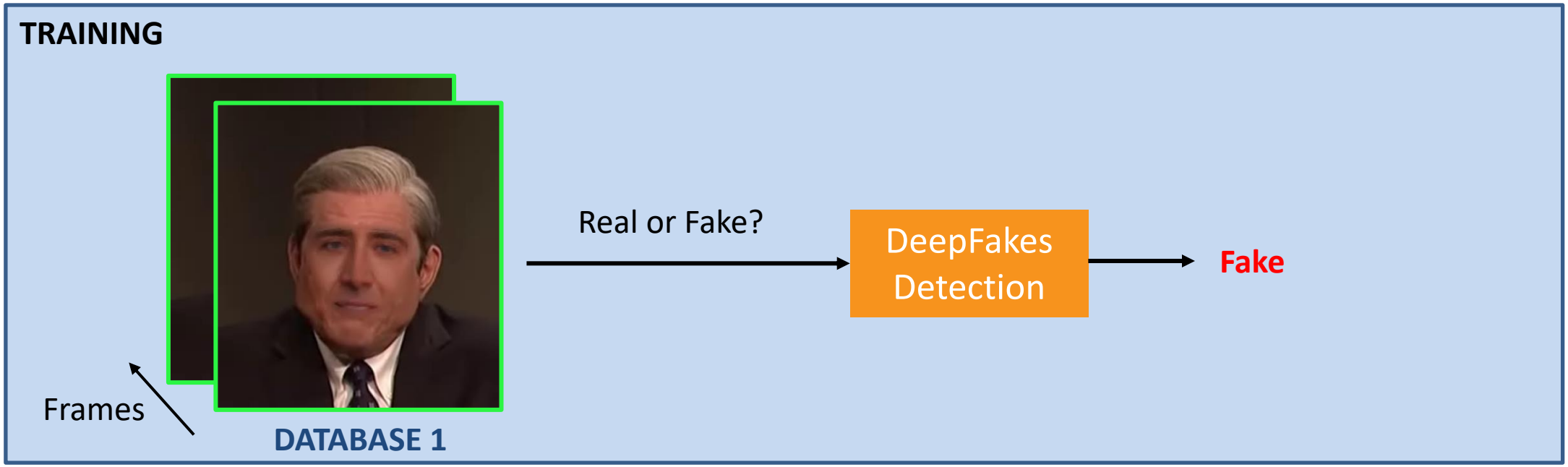
DeepFakes Detection: Inter-Database Analysis

Task 2 - DeepFakes Detection: Inter-Database Analysis

Inter-database analysis: we are going to train and evaluate the systems using different databases.

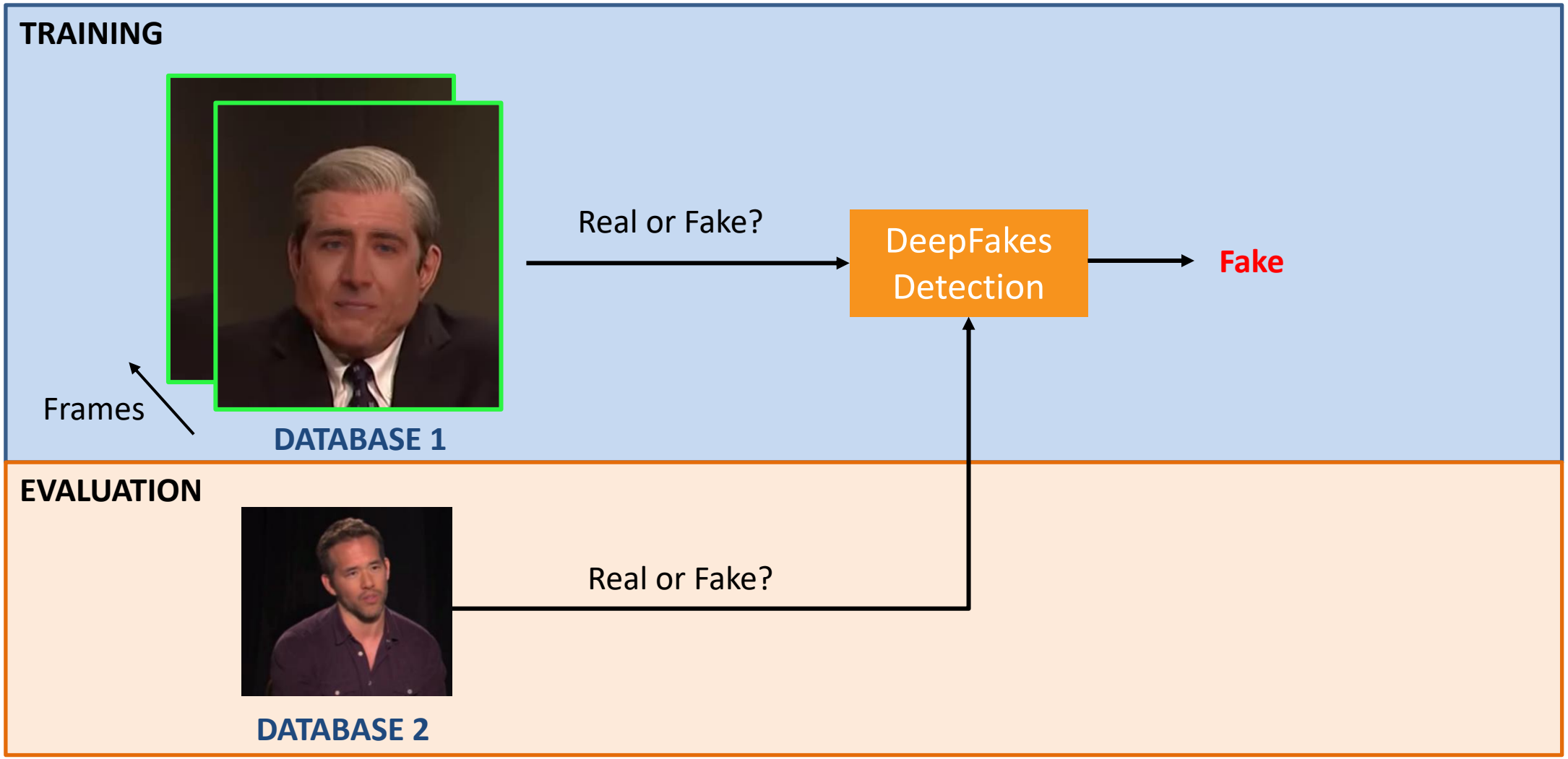
Task 2 - DeepFakes Detection: Inter-Database Analysis

Inter-database analysis: we are going to train and evaluate the systems using different databases.



Task 2 - DeepFakes Detection: Inter-Database Analysis

Inter-database analysis: we are going to train and evaluate the systems using different databases.



Task 2 - DeepFakes Detection: Inter-Database Analysis

Celeb-DF database (v1): 408 real and 795 fake videos using Deep Learning (Autoencoders).

- Y. Li, X. Yang, P. Sun, H. Qi, and S. Lyu, “Celeb-DF: A LargeScale Challenging Dataset for DeepFake Forensics,” in Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 2020.

Task 2 - DeepFakes Detection: Inter-Database Analysis

Experimental protocol:

- Development: the same detection system developed in Task 1 (using only the UADFV database).

- Evaluation: only the Celeb-DF database is considered for the final evaluation of the system (not for training!!!).

  - Evaluation folder:

    - Real folder: 30 videos.

    - Fake folder: 30 videos.

For each video, we provide 10 frames.

TASK 1

DeepFakes Detection:
Intra-Database Analysis

TASK 2

DeepFakes Detection:
Inter-Database Analysis

Restrictions: It is forbidden to use **pre-trained models**.
Only the **development dataset** can be used.

TASK 3

DeepFakes Detection: Inter-Database Proposal

Task 3 - DeepFakes Detection: Inter-Database Analysis

Goal: propose a new DeepFake detection system in order to achieve better inter-database results.

Task 3 - DeepFakes Detection: Inter-Database Analysis

Goal: propose a new DeepFake detection system in order to achieve better inter-database results.

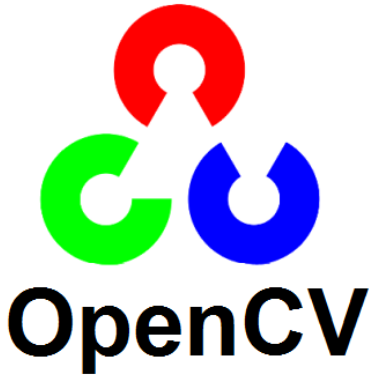
- Development: no restrictions.

- Evaluation: only the Celeb-DF database is considered for the final evaluation of the system (not for training!!!).
  - Evaluation folder:
    - Real folder: 30 videos.
    - Fake folder: 30 videos.

For each video, we provide 10 frames.

Tips

Software and databases: no restrictions.



Evaluation

This Lab is carried out in groups of 3 people.

Evaluation of the Lab:

- Report (max. 8 pages): complete the Word template provided including the information and results of your proposed DeepFake detection system. Do not forget to include all your code when submitting via Moodle.

- Oral presentation (max. 15 minutes): each group should explain the proposal and results achieved in each of the three Tasks.