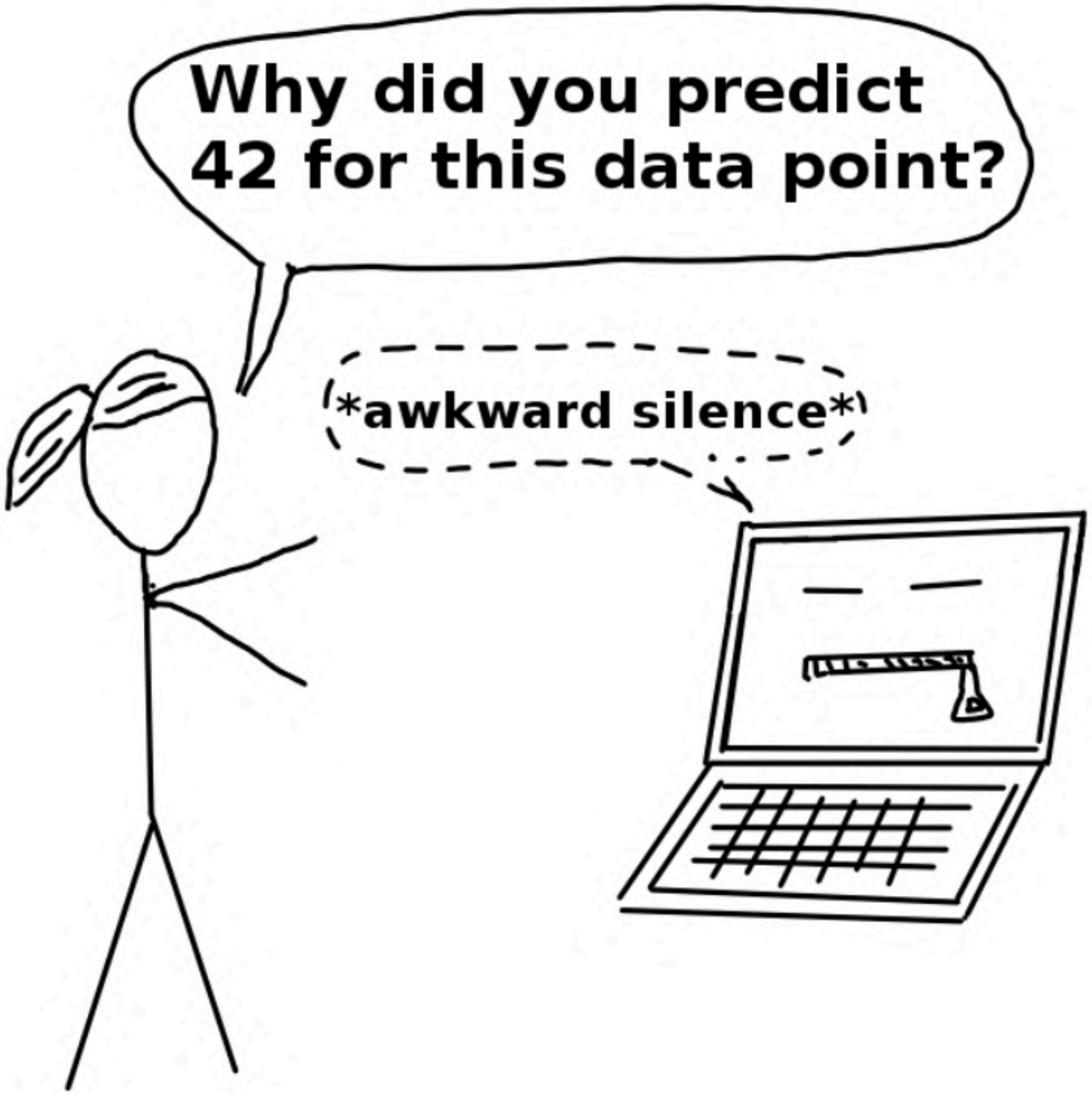
JOÃO BRITO | PROF. DR. HUGO PROENÇA | JANUARY 2021

Deep Adversarial Frameworks for Visually Interpretable Periocular Recognition



[1]

Table of Contents

Problem Description

Problem Description

Given two images from the **periocular region**, determine whether they came from the **same subject** and **explain** why

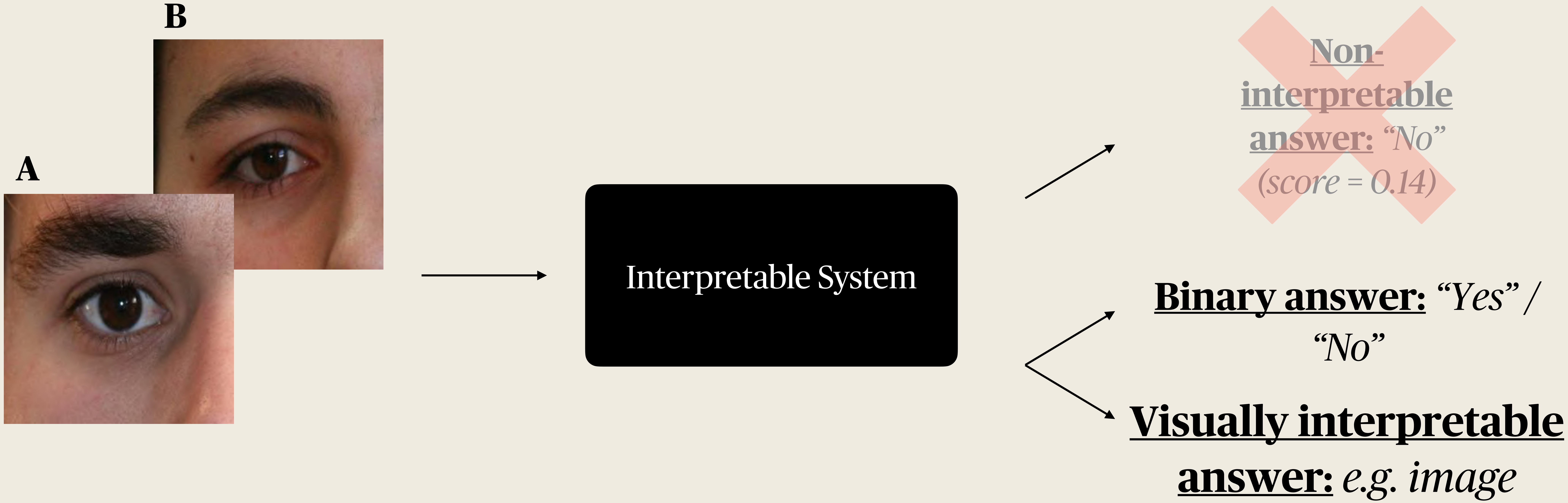


Figure 1: Visualisation of the underlying problem.

Related Work

Related Work

Deep Learning - DenseNet

- **Densely connected** blocks (i.e. within each block, each node receives all the feature maps from the previous nodes)
- **Composition of operations** (e.g. BN, convolutions or pooling)
- Multiple **information pathways**
- Addition of "**transition layers**" in between dense blocks, to change the feature maps' sizes

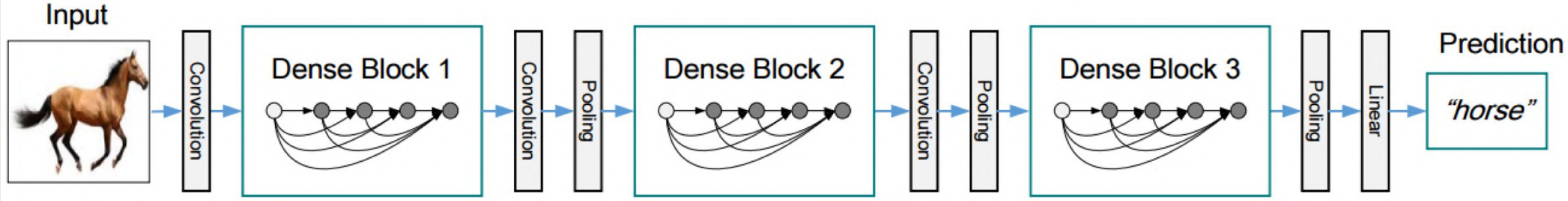


Figure 2: DenseNet architecture with three dense blocks [2].

Related Work

Deep Learning - Mask R-CNN

- Based on the **Faster R-CNN** architecture
- Use of a **RPN** to **propose object regions**
- **Three branches** (classification, regression and segmentation)
- Introduction of **RoIAlign** to **extract the region proposal's feature maps** from the common feature map
- Useful for **instance segmentation** and even **pose estimation**

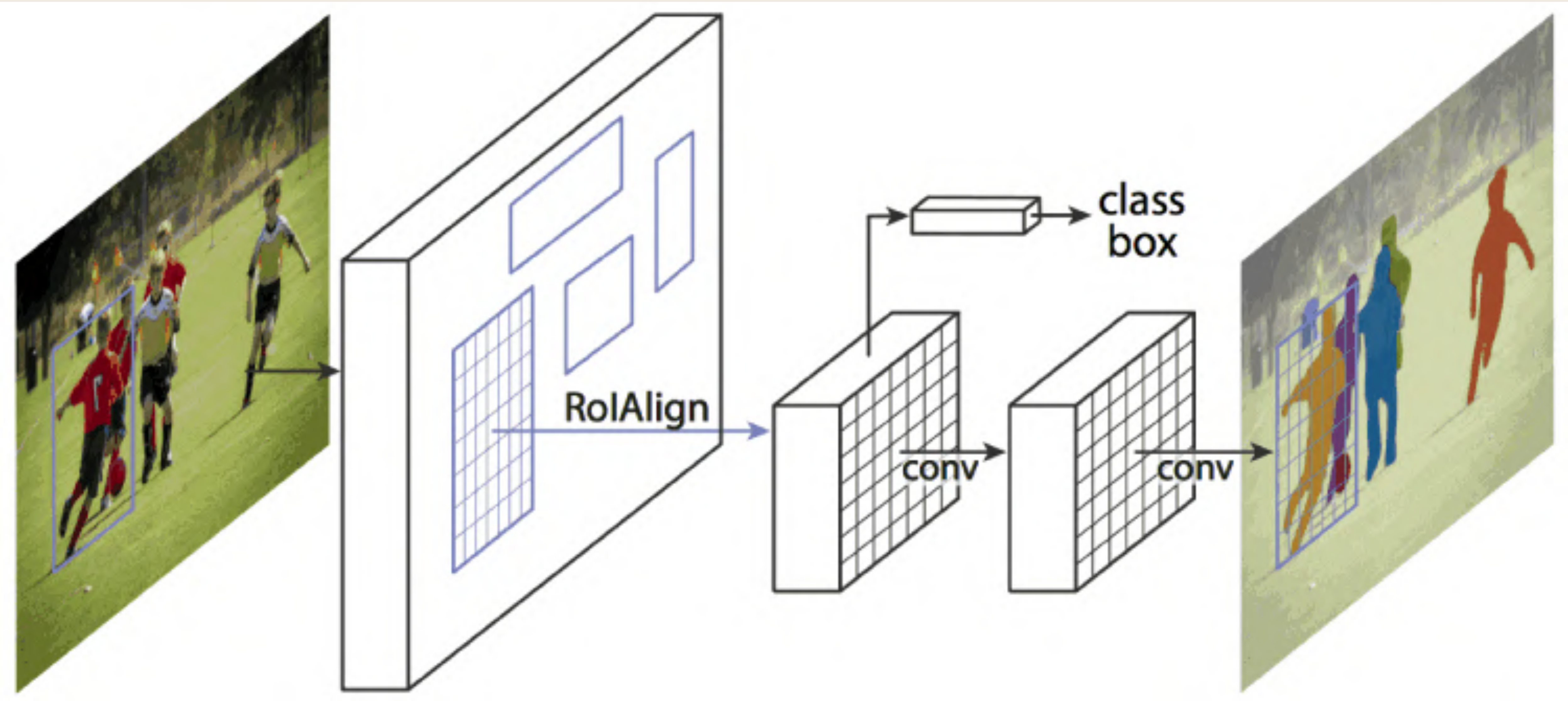


Figure 3: Mask R-CNN architecture, with the small FCN and RoIAlign highlighted [3].

Related Work

Deep Learning - StyleGAN2

- Use of a **mapping network** to map the original latent space (*Z*) to an intermediate, more **disentangled** one (*W*)
- Notion of "**style**", to condition the generation process
- **Incorporation of noise** to add further detail
- **Specifically on StyleGAN2, replacement of progressive growing with input/output skips** to enable the use of multiple resolutions (i.e. keep the advantages of PG) but only if it becomes useful (i.e. thus freeing the networks to use the resolutions as required)

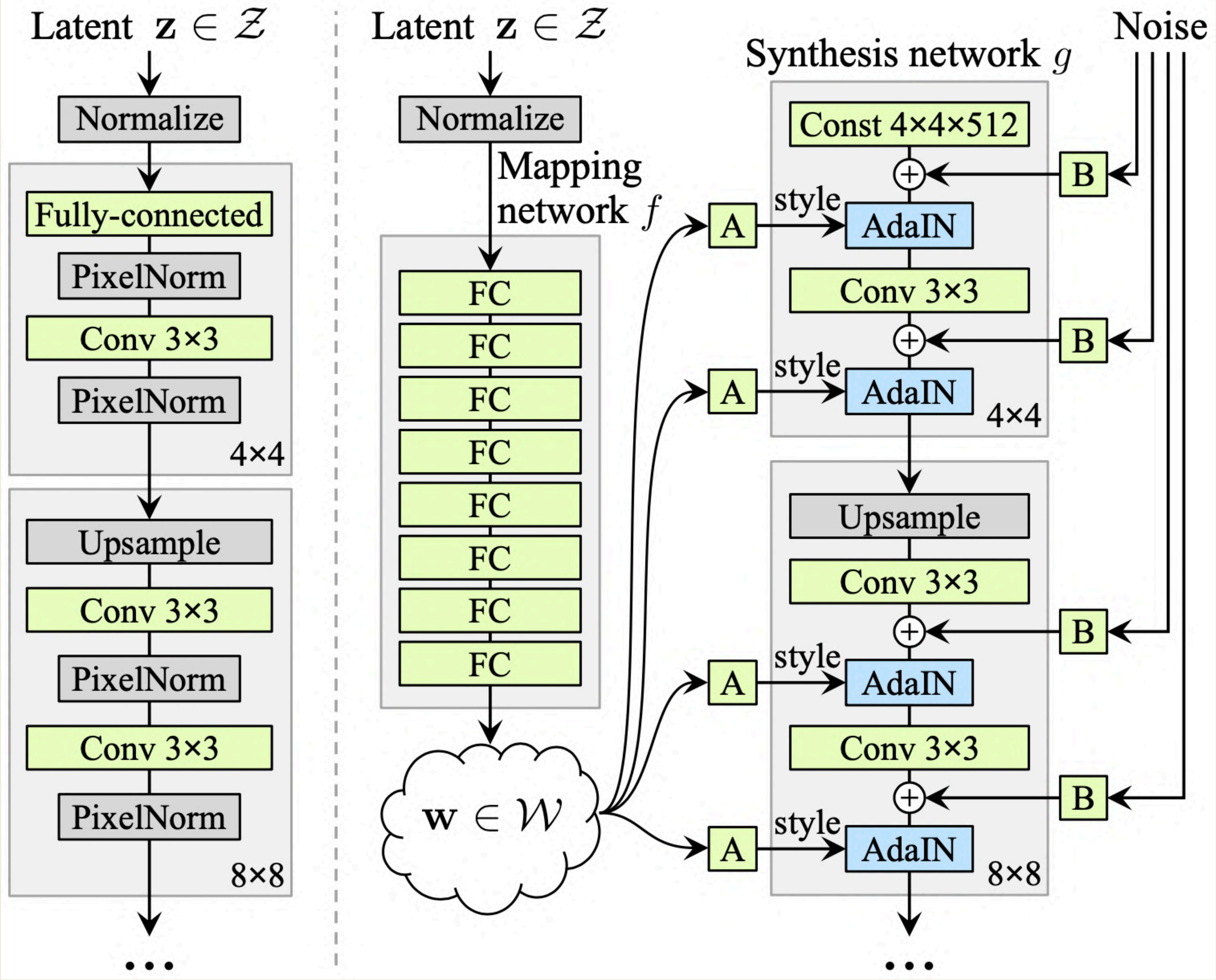


Figure 4: ProgGAN and StyleGAN architectures [4].

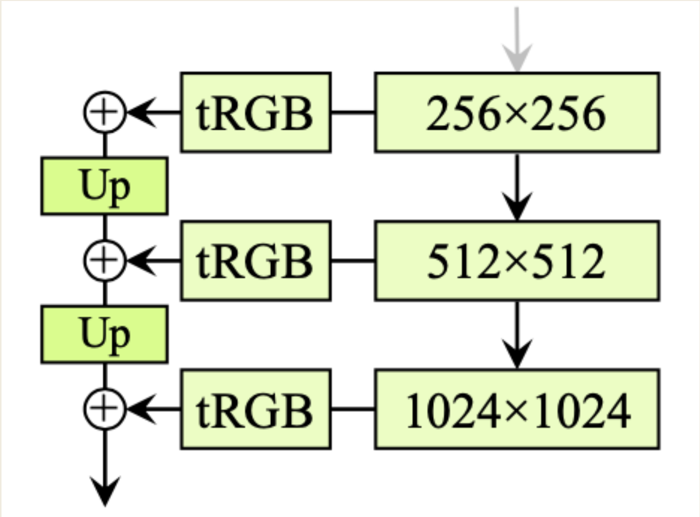


Figure 5: Revised generator architecture for StyleGAN2 [5].

Related Work

ML Interpretability - **L**ocal **I**nterpretable **M**odel agnostic **E**xplanations

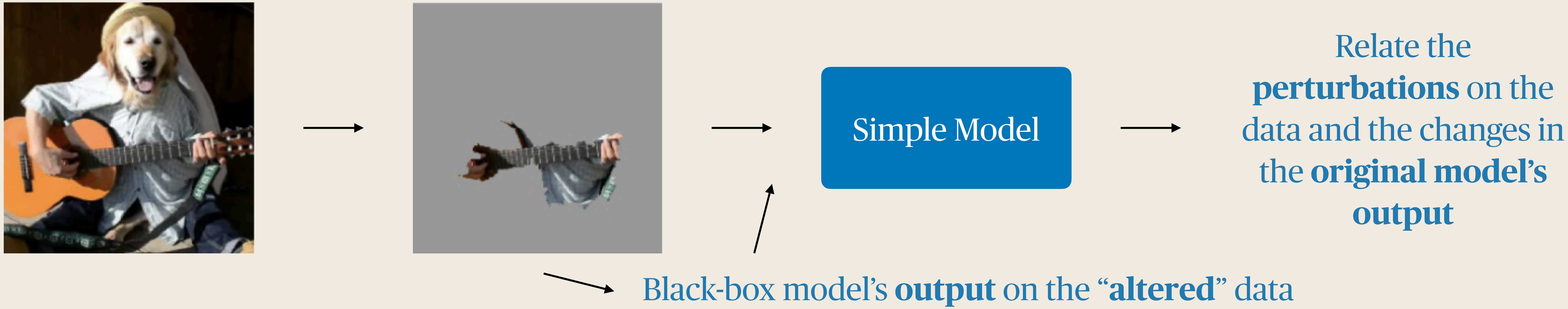


Figure 6: Overview of LIME's explanation process.

$$explanation(x) = \arg\min_{g \in G} \underbrace{L(f, g, \pi_x)}_{\text{Loss}} + \underbrace{\Omega(g)}_{\text{Complexity}}$$

Figure 7: Equation that formally represents an explanation [6].

Related Work

ML Interpretability - Kernel **SH**apley **A**dditive ex**P**lanations

1. Sample ***K*** coalitions: $z'_k \in \{0,1\}^M$
2. Obtain **predictions** for each z'_k by using $f(h_x(z'_k))$
3. Compute the **weight** for each z'_k with equation 2.20
4. Fit the **weighted linear model** using equation 2.18
5. Return the **coefficients** from the linear model, namely, the **Shapley values** ϕ_i

$$\pi_x(z') = \frac{(M-1)}{\binom{M}{|z'|}|z'|(M-|z'|)} \tag{2.20}$$

$$L(f,g,\pi_x) = \sum_{z' \in Z} [f(h_x(z')) - g(z')]^2 \pi_x(z') \tag{2.18}$$

Figure 8: The top equation defines the SHAP Kernel, which gives more or less weight to small / large coalitions. The bottom equation defines the loss function when using KernelSHAP [7].

Approach with LIME/SHAP

Approach with LIME/SHAP

1. Traditional **CNN-based classification** of the test pair **(G**enuine or **I**mpostor)
2. **Prediction interpretation** using interpretability techniques (e.g. **LIME**, **KernelSHAP**, etc...)

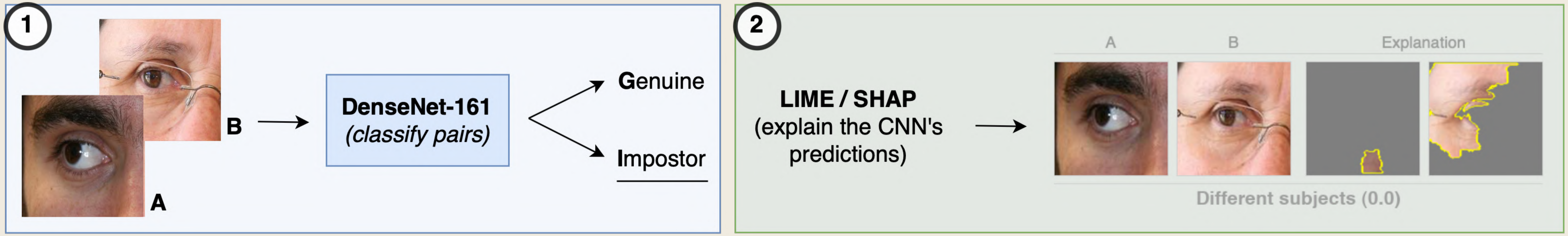


Figure 9: Results obtained with LIME / SHAP on top of a traditional CNN.

Approach with LIME/SHAP

Figure 10: Genuine pairs correctly classified by the CNN.

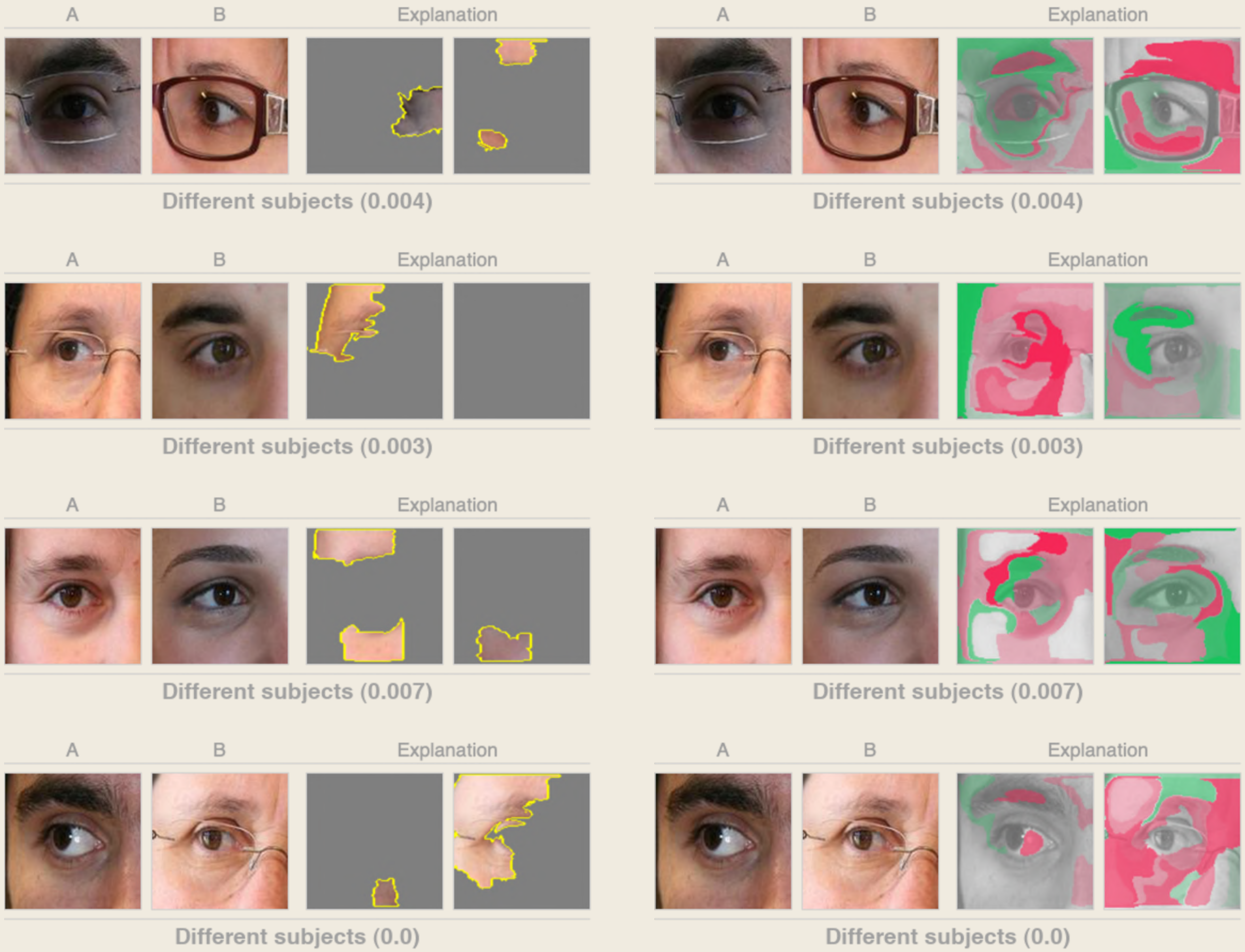


Figure 11: Impostor pairs explained with LIME (on the left) and KernelSHAP (on the right).

Proposed Method

Proposed Method

1. Traditional **CNN-based classification** of the test pair (**G**enuine or **I**mpostor)
2. **Segmentation** of iris, sclera and eyebrow
3. **Synthetic pair** generation and **segmentation**
4. **Comparison and gathering** of the ***K*** closest synthetic pairs (i.e. neighbours)
5. **Pixel difference** between the test image ***B*** and the neighbours' images ***B***

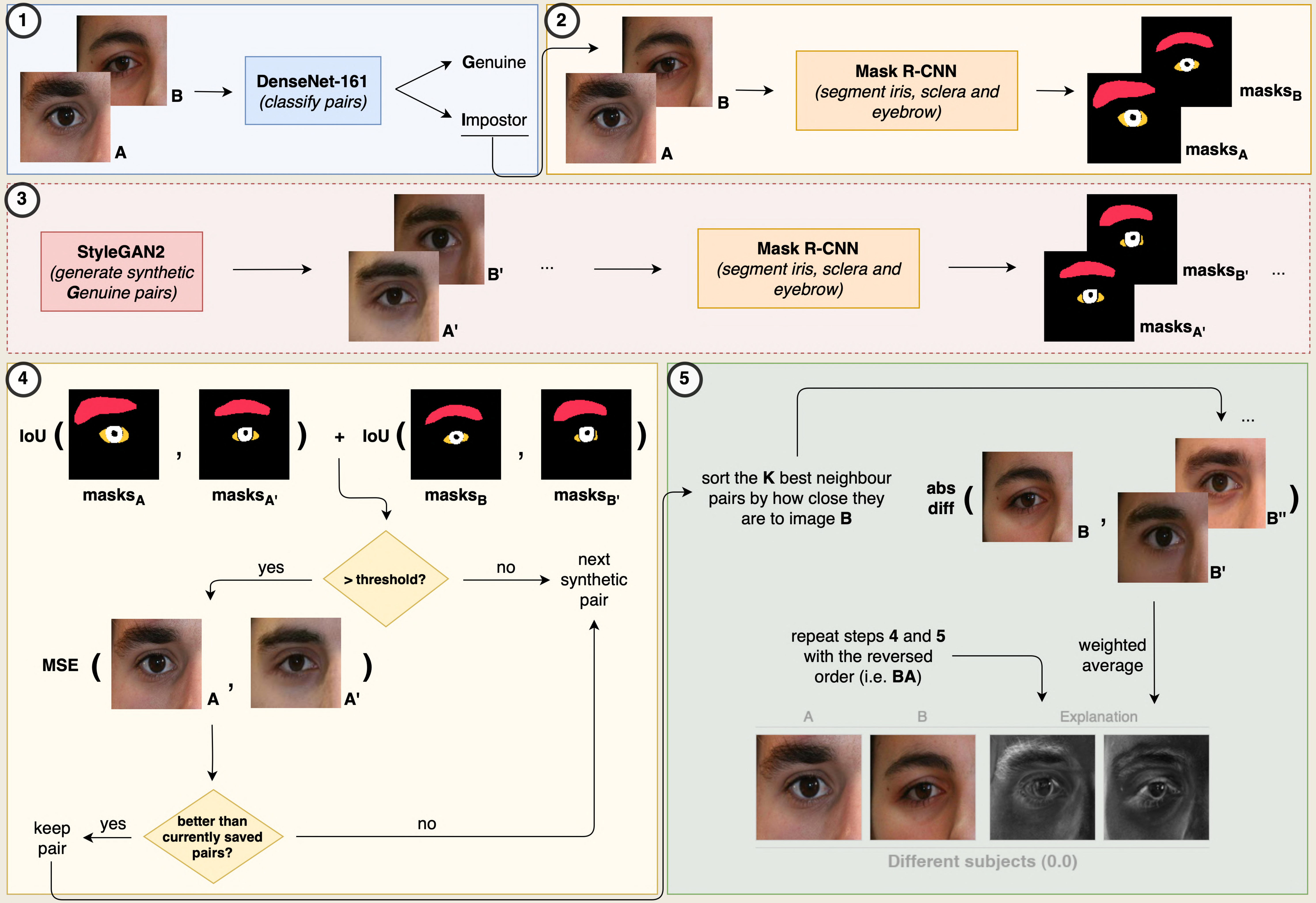


Figure 12: Diagram of the main pipeline.

Proposed Method

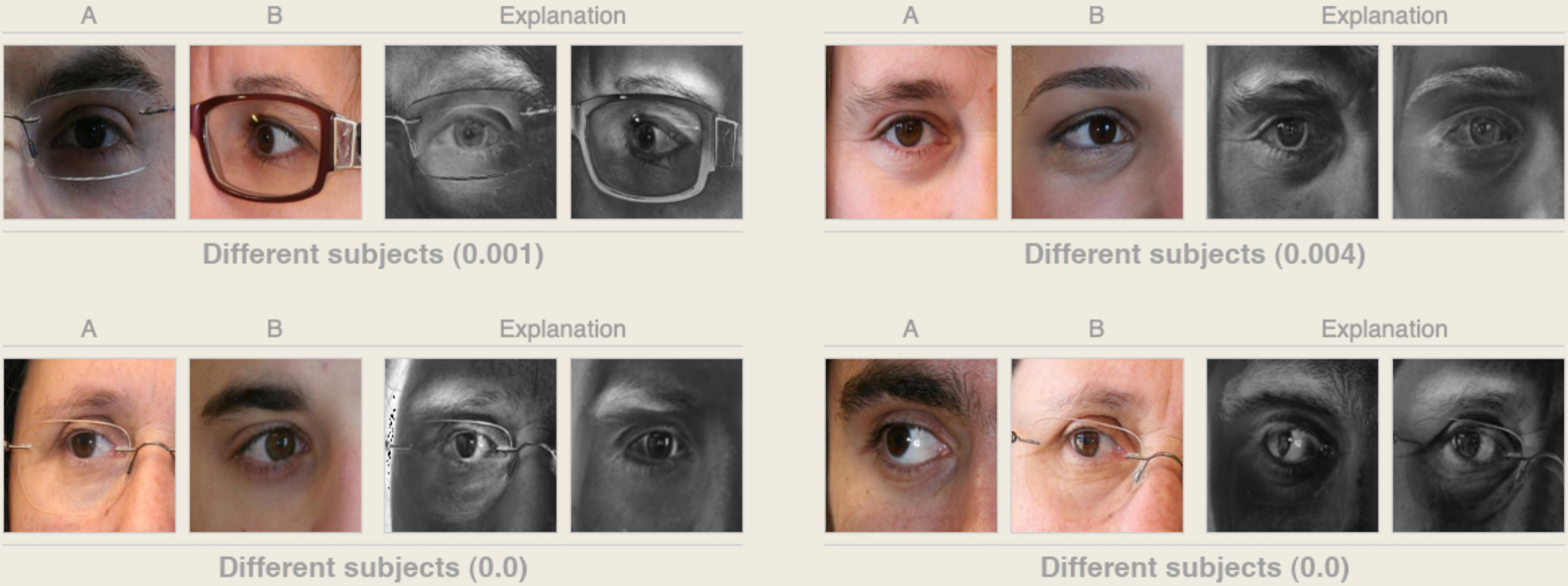


Figure 13: Results obtained with the first method (directly comparable to LIME/SHAP).

Proposed Method

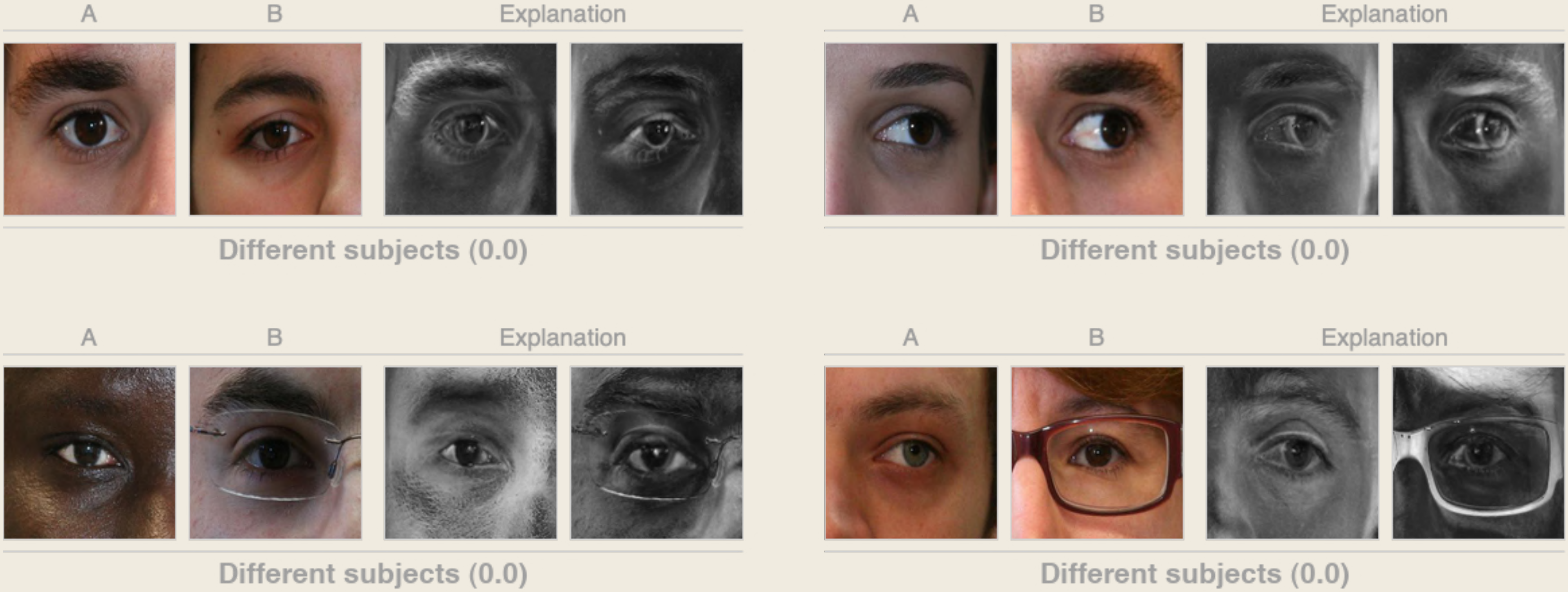


Figure 14: Results obtained with the first method.

Ongoing and Future Tasks

Ongoing and Future Tasks

- Use **other techniques** besides IoU
- Add **captions** to the visual explanations (i.e. a text representation)
- Make **speed** improvements
- Consider **other approaches** (possibly based on **latent space directions**)

Conclusion

Any questions?

Bibliography

1. Christoph Molnar. Interpretable Machine Learning. 2019. Available from: https://christophm.github.io/interpretablemlbook/
2. Laurens van der Maaten Gao Huang, Zhuang Liu and Kilian Q. Weinberger. Densely connected convolutional networks, 2018
3. Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation, 2015
4. Tero Karras, Samuli Laine, and Timo Aila. A style based generator architecture for generative adversarial networks, 2019
5. Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan, 2020
6. Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?": Explaining the predictions of any classifier, 2016
7. Scott Lundberg and SuIn Lee. A unified approach to interpreting model predictions, 2017
8. USA Home Listings. Available from: https://www.usahomelistings.com/blog/moving-forward-marketing-trends-to-watch-out-for-in-2020/