Summary of Annotation Task

We designed an annotation task to address potential biases (gender, age, and ethnicity) in the AffectNet dataset, frequently referenced in facial expression recognition (FER) research. The project aims to assess these biases by combining automated and manual annotation approaches. Initially, we employed the DeepFace Python library for automated labeling of images, but to gauge the accuracy and reliability of this process, we are conducting manual annotations as well.

Given the vast size of AffectNet, we are still determining the appropriate volume of manual annotations needed to achieve reliable metric calculations, such as accuracy, consistency, and inter-annotator agreement. We plan to use metrics like Cohen's Kappa to compare manual annotations among the different 4 annotators and against the automated results. To establish a baseline and explore the available techniques, we created an example set of 12 images, ensuring diverse representation across age groups, ethnicities, and an equal number of male and female subjects. As seen in the table on the next page.

Manual and Automatic Annotation Categories

To ensure comparability between manual and automated annotations, we aligned the categories for manual labeling with those used by DeepFace:

Gender: DeepFace typically returns "Man" or "Woman" for gender, and we used the same categories for manual annotations.

Race/Ethnicity: DeepFace assigns a dominant race from the following categories:
- Asian
- Indian
- Black
- White
- Middle Eastern
- Latino Hispanic

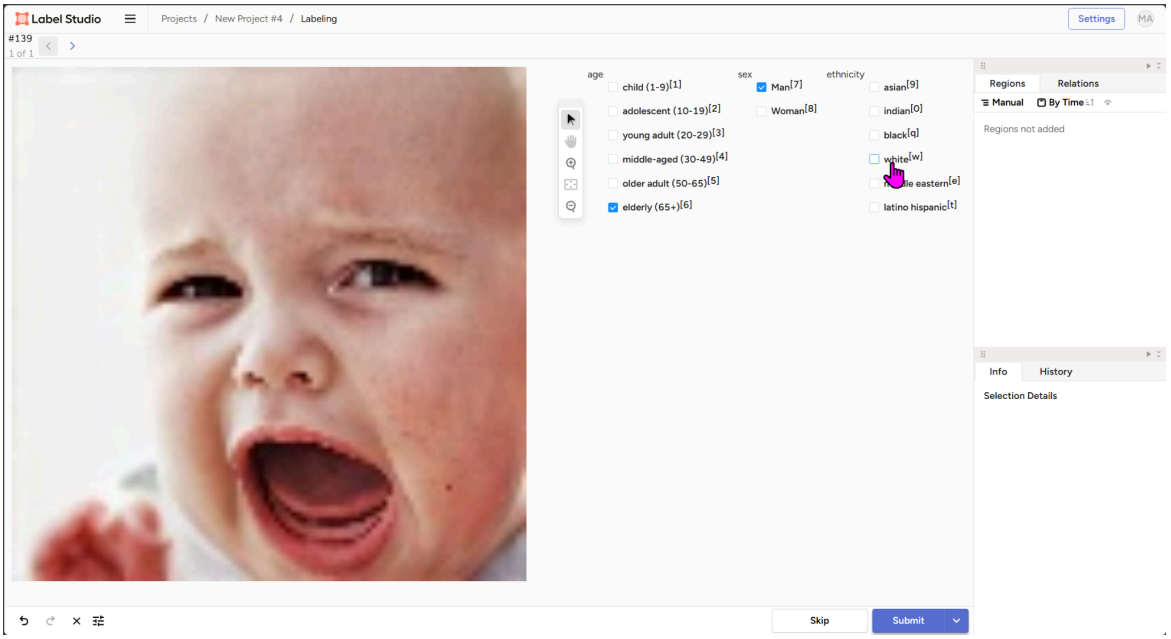
Age: DeepFace provides age as an integer. For consistency and ease of manual annotation, we convert these ages into age groups: baby (0-1), child (1-9), adolescent (10-19), young adult (20-29), middle-aged (30-49), older adult (50-65), and elderly (65+).

Selected Pictures

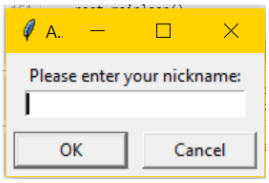
age/ethnicity	Asian	Indian	Black	White	Middle eastern	Latino hispanic
baby (0-1)				Image0001063 anger 		
child (1-9)	Ffhq_130 neutral 					Ffhq_7 happy 
adolescent (10-19)		Image0001685 fear 	Image0000866 Disgust 			
young adult (20-29)		Ffhq_15 neutral 				Image0003731 Fear 
middle-aged (30-49)			Ffhq_164 happy 		Image0003108 Anger 	
older adult (50-65)				Ffhq_107 neutral 		
and elderly (65+)	Image0013196 Contempt 				Image0006237 sad 	

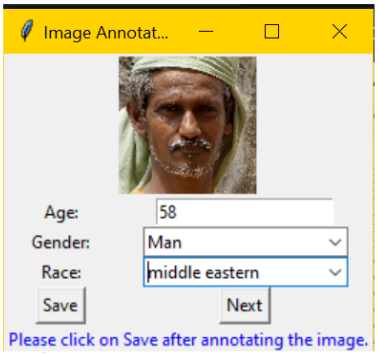
Interfaces

1. We uploaded this example set to Label Studio to facilitate manual annotations, as shown in the picture.



2. Additionally, we designed a manual annotation interface that can run as a Python script with pop-up windows, allowing for a more interactive annotation process. Results are saved in a separate CSV file for each annotator.





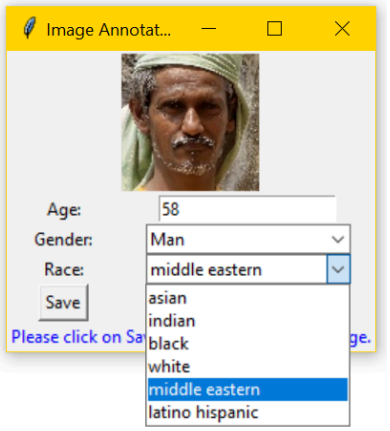


	Image	Age	Gender	Race	Annotation_by
0	image0021846.jpg	65.0	Man	middle eastern	Idk
1	image0021851.jpg	56.0	Woman	white	Idk
2	image0021894.jpg	29.0	Woman	white	Idk
3	image0021941.jpg	39.0	Man	white	Idk
4	image0021869.jpg	65.0	man	latino hispanic	Idk
5	image0021955.jpg	45.0	Man	indian	Idk

The choice of which design to use will depend on the ease of sharing.

Annotation Guide Overview

We have created 2 annotation guides for the 2 different approaches described above.

Annotation Guide for Label Studio

Summary:

The annotation task utilizes Label Studio software to facilitate the labeling process. Please follow the detailed instructions provided in the README PDF file within the zip folder.

You will be required to annotate each image by providing answers for three categories: age, sex, and ethnicity. We have chosen to use age groups (or "buckets") instead of exact ages to make it easier for annotators to quickly assign labels, leading to more spontaneous and intuitive data.

After completing the annotations, please follow the exporting instructions outlined in the PDF.

Guidelines:

Do’s:

- Provide exactly one answer for each category (age, sex, ethnicity) for every image.
- Ensure that your annotations are as accurate as possible based on your judgment.

Don'ts:

- Do not worry about making incorrect guesses. Part of the analysis is to identify which images are more challenging to label, so your honest attempt is valuable.
- Avoid leaving any categories blank.

Additional Notes:

- If you are unsure about how to categorize an image, refer to the examples provided in the README or ask for clarification.
- Remember, your responses are essential for understanding annotation consistency and assessing the performance of automated systems.

Annotation Guide for the Python Script Interface

You will be required to annotate 12 images by providing answers for three categories: age, sex, and ethnicity.

1. Enter a Nickname: Please start by entering a nickname. Choose something that cannot be linked to your identity.
2. Annotate Each Image: You will then see the first image. Select an option for each category (age, sex, ethnicity) from the drop-down menus. Ensure that you provide exactly one answer for each category for every image, based on your best judgment.
3. Save Your Selections: After you have made selections for all three categories, click on "Save."
4. Proceed to the Next Image: Click "Next" to move on to the next image.
5. Repeat the Process: Continue this process for all images.
6. Review Completed Annotations: After completing all annotations, please review your entries to ensure that no fields have been left empty.

Do not worry about making incorrect guesses. Part of the analysis is to identify which images are more challenging to label, so your honest effort is valuable.

Remember, your responses are essential for understanding annotation consistency and assessing the performance of automated systems.