

Appendix of the paper **Semiotic-Based Construction of a Large Emotional Image Dataset with Neutral Samples**

Marco Blanchini, Giovanna Dimitri, Lydia Abady,
Benedetta Tondi, Tarcisio Lancioni, Mauro Barni
University of Siena, Italy

m.blanchini@student.unisi.it

1. Introduction

Hereby we will provide additional information and examples of images included in our paper *Semiotic-Based Construction of a Large Emotional Image Dataset with Neutral Samples*. The methodology for building the dataset was described in Section 3 of the main paper. The full list of subjects, adjectives, and noun forms of adjectives used to create our dataset is described in Table 1. The appendix is organized as follows: In Section 2, we provide further details on the dataset construction, including the complete list of queries used. In Section 3, we will provide examples of images taken from our dataset of the three categories: neutral, positive, and negative. In Section 4, we will further specify and address issues related to the usability of the images contained in the dataset, with specific reference to images involving individuals. In Section 5, we provide the link to download the dataset.

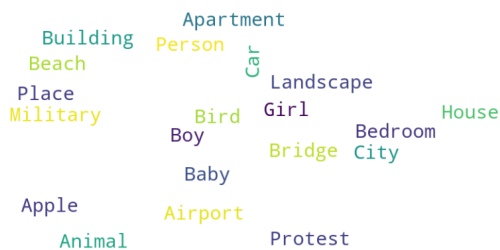


Figure 1. Cloud representation of the list of subjects used to form the queries to build our dataset

2. Supplementary Details on Dataset Construction

In this section, we provide further details on the construction of the dataset and the queries used. The subjects selected for the construction of the dataset are: Airport, An-

imal, Apple, Apartment, Baby, Beach, Bedroom, Bird, Boy, Bridge, Building, Car, City, Girl, House, Landscape, Military, Person, Place, Protest, as represented in Figure 1. In selecting the subjects, we aimed to include both subjects that are visually very similar and tend to have many traits in common, such as *Bird* and *Bridge*, as well as subjects with a wider range of visual forms, such as *Animal* and *Place*. All the subjects we selected are real subjects, allowing for a tangible visual representation. The list of selected subjects is shown in Figure 1. Starting from these subjects, we selected pairs of adjectives with opposite emotional values that are frequently associated with them. Subsequently, using the methodology described in Section 3 of the paper, we constructed the queries to identify the images. Table 1 shows the complete list of queries employed.

In the first column, we list the pairs of adjectives with opposing emotional values used in the queries. The adjective pairs were identified using the NLTK [4] (Natural Language Toolkit) library to search for the adjectives most commonly associated with the selected subjects in the English language. This library allowed us to create a code to identify the 100 words most frequently associated with the subjects. Among these, we selected the adjectives with the clearest emotional value. To determine which of the identified adjectives had a clearer emotional content in terms of positivity and negativity, we used the SentiWordNet textual sentiment analysis tool [5]. Once the adjectives with the most distinct emotional value were identified, we found their antonyms, that is, adjectives with the opposite meaning. We then paired the adjectives with their antonyms and created a list of adjective pairs for the queries. Each pair of adjectives was then used to form the queries with multiple nouns. In some cases, instead of the antonym of an adjective as indicated by the antonym dictionary [6], we preferred a near-antonym, which is an adjective with a nearly opposite meaning. For example, in the case of the adjective “Dead”, the antonym would be the adjective “Alive”. However, “Alive” does not have a distinctly positive value, so we preferred the near-antonym “Vital”, which has a markedly

positive value. The antonyms and near-antonyms of the adjectives were identified using the Merriam-Webster online dictionary [6]. The evaluation of the emotional value of antonymic and near-antonymic adjectives was carried out using SentiWordNet [5].

The second and third columns contain the nominalized adjectives derived from the positive and negative adjectives, which were used to download artistic images representing these concepts. In the fourth column, we indicate the neutral adjectives identified for each opposition of emotionally relevant adjectives associated with a subject’s name. In some cases, the neutral adjectives identified using the semiotic square can be suitable for different pairs of opposite adjectives related to the same subject. In these cases, multiple images were downloaded for the combination “neutral adjective + subject”, depending on the number of pairs of opposite adjectives for which the neutral adjective was appropriate. For example, the adjective “Normal” is suitable both for the association between “Perfect - Damaged” with the subject “Apple” and for the association between “Fresh - Rotten” with the same subject “Apple”. In this case, additional images were downloaded for the query “Normal + Apple”. In the fifth column, we specify the subjects we paired with the adjective pairs to form the queries. Each pair of adjectives was associated with a number between 3 and 7 subjects. If a sufficient number of images was not available for some adjective-subject pairs, we supplemented the search by replacing the adjectives with synonyms. The search for synonyms was conducted using the Merriam-Webster Thesaurus [6], the online dictionary of synonyms and antonyms, selecting the closest synonyms of the adjectives.

3. Our dataset examples

In this section, we present two examples of how we constructed the dataset, as described in Section 3 of the main paper. Specifically, we demonstrate how we used the semiotic square to form the queries and provide examples of the images identified using the selected queries. Figure 2 contains a flowchart that illustrates the procedure for creating the queries. The query formation procedure is as follows: A subject name was selected from the list, for example, “House”.

A pair of opposing adjectives suitable for the subject was identified, for example, “Luxurious” and “Shabby”.

The adjectives were transformed into nouns. These nouns are not associated with the subject “House” but are used as standalone queries to identify images associated with the artistic representation of the concept in general. For example, “Shabby” into “Shabbiness” and “Luxurious” into “Luxury”.

The semiotic square was applied to the pair of opposing adjectives combined with the subject to identify an appropri-

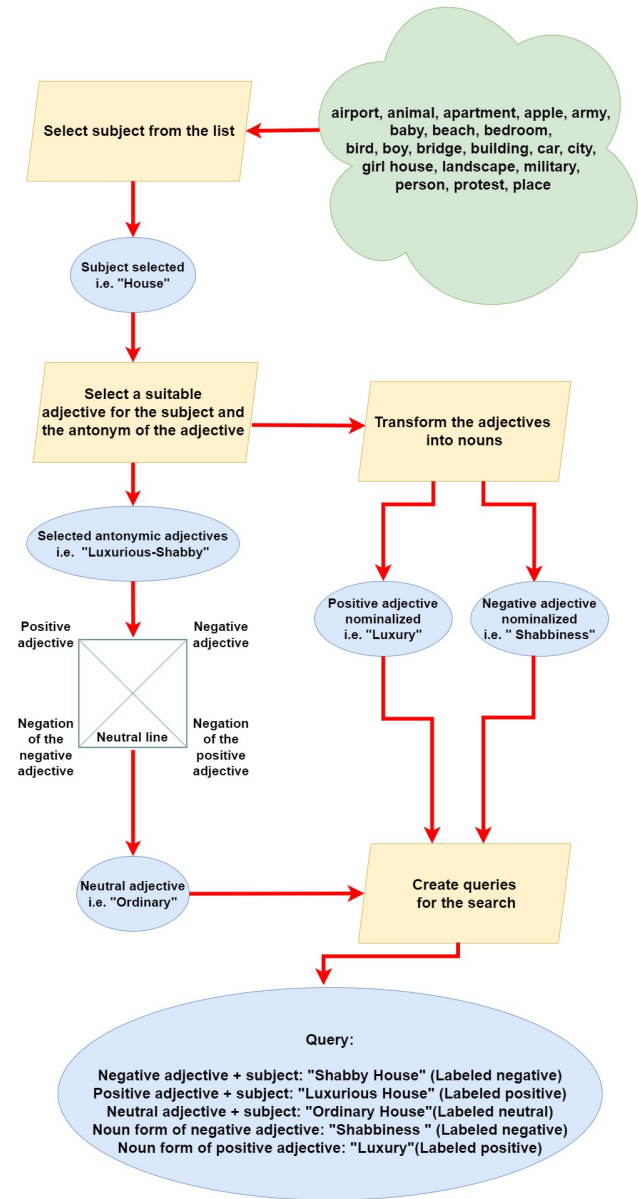


Figure 2. Flowchart of the application of the semiotic square in the query selection for our dataset

ate neutral adjective, for example, the adjective “Ordinary” which is neutral with respect to the opposition “Luxurious” and “Shabby”.

The five queries were formed as follows:

Negative adjective + subject: “Shabby House”

Positive adjective + subject: “Luxurious House”

Neutral adjective + subject: “Ordinary House”

Noun form of negative adjective: “Shabbiness”

Noun form of positive adjective: “Luxury”

The images identified with the queries subject + positive

OPPOSITE ADJECTIVES	POSITIVE NOUNS	NEGATIVE NOUNS	NEUTRAL ADJECTIVES	SUBJECTS MATCHED
Admirable - Deplorable	Admirability	Deplorability	Average, Common, Neutral	Boy, Girl, Military, Person
Adorable - Loathsome	Adorability	Loathing	Common, Normal, Unremarkable	Animal, Baby, Bird, Boy, Girl
Attractive - Unattractive	Attractiveness	Unattractiveness	Common, Ordinary, Unremarkable	City, Person, Place
Beautiful - Ugly	Beauty	Ugliness	Average, Ordinary, Unremarkable	Airport, Animal, Bird, Car, City, House, Landscape
Bright - Dismal	Brightness	Dismality	Common, Ordinary, Regular	Airport, Beach, Bridge, Building, House, Person
Cheerful - Depressed	Cheerfulness	Depression	Indifferent, Unsentimental	Baby, Boy, Girl, Person
Clean - Dirty	Cleanliness	Dirtyness	Average, Ordinary, Regular, Standard	Apartment, Bedroom, Building, Car, City, House, Place
Comforting - Terrifying	Comfort	Terror	Common, Neutral, Standard	Bedroom, Building, Car, House, Place
Cozy - Unhomely	Coziness	Unhomeliness	Average, Ordinary, Regular	Apartment, Beach, Bedroom, Car, House
Delighted - Angry	Delighted	Anger	Indifferent, Neutral, Unsentimental	Animal, Baby, Boy, Girl, Person
Delicious - Horrid	Deliciousness	Horror	Normal, Ordinary, Regular, Unremarkable	Animal, Apple, Bird, House, Person, Place
Enthusiastic - Bored	Enthusiasm	Boredom	Indifferent, Unsentimental	Baby, Boy, Girl, Person
Friendly - Hostile	Friendliness	Hostility	Indifferent, Neutral	Animal, Boy, Girl, Military, Person
Fresh - Rotten	Freshness	Rottenness	Normal, Neutral, Ordinary	Apple, Bedroom, House, Landscape, Place
Good - Bad	Goodness	Badness	Neutral	Animal, Boy, Girl, Person
Healthy - Sick	Health	Sickness	Neutral, Normal, Regular	Animal, Baby, Boy, Girl, Person
Joyous - Melancholic	Joy	Melancholy	Indifferent, Neutral, Unsentimental	Baby, Boy, Girl, House, Landscape, Person
Kind - Unsympathetic	Kindness	Unsympathy	Indifferent, Neutral	Animal, Baby, Boy, Girl, Person
Lovable - Hateful	Love	Hate	Unremarkable	Animal, Bird, Boy, Girl, Person
Luxurious - Shabby	Luxury	Shabbiness	Normal, Ordinary, Standard	Airport, Apartment, Building, Car, House, Place
Mild - Ferocious	Mildness	Ferocity	Neutral	Animal, Bird, Boy, Person
Ordered - Chaotic	Order	Chaos	Average, Common, Regular	Airport, Apartment, Beach, Building, House
Peaceful - Violent	Peace	Violence	Calm	Military, Person, Protest
Pleased - Sorrowful	Pleasure	Sorrow	Indifferent, Unsentimental	Animal, Baby, Boy, Girl, Person
Pleasing - Unpleasant	Pleasure	Unpleasantness	Average, Common, Neutral	Baby, Beach, Boy, Girl, Person
Pure - Polluted	Purity	Pollution	Average, Common, Ordinary	Beach, Landscape, Place
Reassured - Worried	Reassurance	Worry	Indifferent	Baby, Boy, Girl, Person
Relaxed - Stressed	Relaxation	Stress	Indifferent, Neutral, Normal	Boy, Girl, Person, Place
Safe - Dangerous	Safety	Danger	Regular, Standard	Bridge, Building, Car, House, Place, Protest
Serene - Distressed	Serenity	Distress	Imperturbable, Indifferent, Unsentimental	Animal, Baby, Boy, Girl, Person
Smiling - Tearful	Smile	Tears	Indifferent, Unsentimental, Neutral	Baby, Boy, Girl, Person
Strong - Weak	Strength	Weakness	Average, Regular	Boy, Girl, Military, Person, Protest
Sweet - Rude	Sweetness	Rudeness	Indifferent	Baby, Boy, Girl, Person
Tidy - Messy	Tidiness	Messiness	Normal, Ordinary, Regular	Airport, Apartment, Bedroom, Building, House, Place
Victorious - Defeated	Victory	Defeat	Neutral	Military, Person, Protest
Vital - Dying	Vitality	Death	Calm, Normal, Ordinary	Bird, Boy, Girl, Person, Place
Wonderful - Awful	Wonder	Awfulness	Average, Common, Ordinary, Regular	Airport, Animal, Bird, Bridge, Building, House, Landscape

Table 1. List of Opposite Adjectives, Corresponding Nouns, Neutral Adjectives and Subjects

adjective and positive adjective nominalized are labeled as positive. The images identified with the queries subject + negative adjective and negative adjective nominalized are labeled as negative. The Images retrieved from the query subject + neutral adjective are labeled as neutral. Each of the examples shows the formation of the queries starting from a subject associated with two opposing adjectives, and the identification of the neutral adjective using the semiotic square. Afterward, the five resulting queries are listed. After each of the two examples, samples of images downloaded for each created query are shown in sections 3.1, 3.2, and 3.3.

3.1. Example I (Apple)

In this first example, we apply the semiotic square to the pair of opposing adjectives “Fresh” and “Rotten” associated with the subject “Apple”. This allows us to identify the term that best expresses the absence of these two distinctive characteristics in the representation of an apple, in this case, “Normal”. We then transform the two adjectives into subjects, “Fresh” into “Freshness” and “Rotten” into “Rottenness”. These two subjects derived from the adjectives are then used to identify artistic images that represent these concepts.

Subject: Apple

Opposite adjective: Fresh VS Rotten

Keyword for the artistic images: Freshness, Rottenness

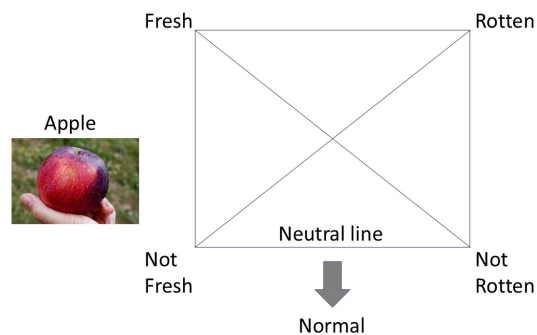


Figure 3. Example I: Example of applying the semiotic square to the pair of adjectives “fresh” and “rotten” paired with the subject “apple”.

Semiotic square representation: Figure 3

Neutral adjective: Normal

List of queries:

- Query positive adjective + subject: Fresh + Apple (labeled positively), see Figure 4
- Query negative adjective + subject: Rotten + Apple (labeled negative), see Figure 5
- Query neutral adjective + subject: Normal + Apple (labeled neutral), see Figure 6



Figure 4. Images retrieved from the query: Fresh + Apple

- Query for positive artistic images: Freshness (labeled positively), see Figure 7
- Query for negative artistic images: Rottenness (labeled negative), see Figure 8

3.2. Example II (Animal)

In this example, we apply the semiotic square to the opposition of the adjectives “Mildness” and “Ferocious” applied to the subject “Animal”. The application of the square allows us to identify the term that best indicates the absence of both of these characteristics in an animal. We then proceed to transform the adjectives into the subjects “Mildness” and “Ferocity” to identify the artistic representations of these concepts.

Subject: Animal

Opposite adjective: Mild VS Ferocious

Keyword for the artistic images: Mildness, Ferocity

Semiotic square application: Figure 9

Neutral adjective: Neutral

List of queries:

- Query positive adjective + subject: Mild + Animal (labeled positive), see Figure 10
- Query negative adjective + subject: Ferocious + Animal (labeled negative), see Figure 11
- Query neutral adjective + subject: Neutral + Animal (labeled neutral), see Figure 12



Figure 5. Images retrieved from the query: Rotten + Apple



Figure 6. Images retrieved from the query: Normal + Apple

- Query for positive artistic images: Mildness (labeled positive), see Figure 13
- Query for negative artistic images: Ferocity (labeled negative), see Figure 14

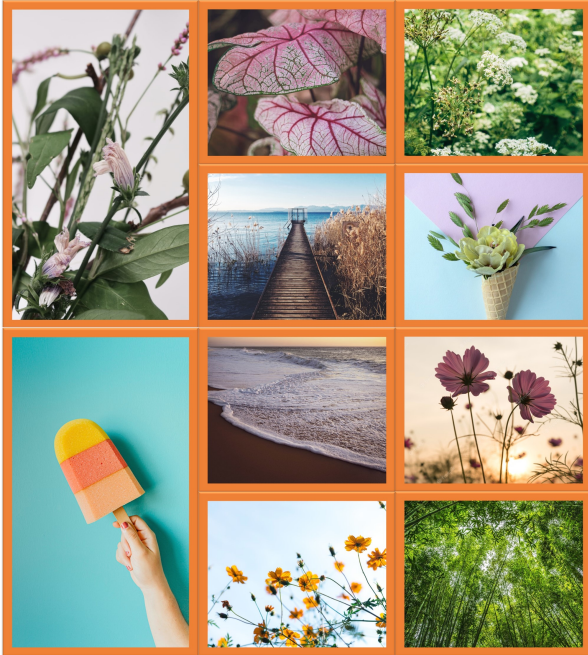


Figure 7. Images retrieved from the query: Freshness



Figure 8. Images retrieved from the query: Rotteness

3.3. Example III (Beach)

In this example, we apply the semiotic square to the opposition of the adjectives “Pure” and “Polluted” applied to the subject “Beach”. In this case, the semiotic square allows us to identify the term that best distinguishes a beach

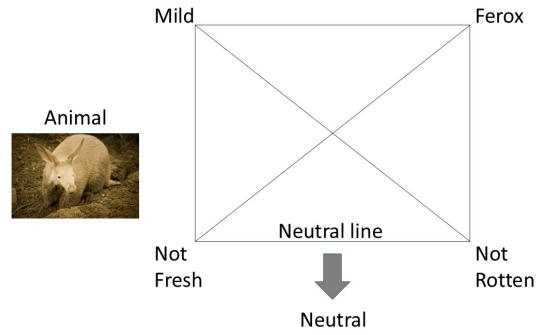


Figure 9. Example II: Semiotic square application



Figure 10. Images retrieved from the query: Mild + Animal

associated with an adjective that lies midway between these two opposites, in this case “Ordinary”.

Subject: Beach

Opposite adjective: Pure VS Polluted

Keyword for the artistic images: Purity, Pollution

Semiotic square application: Figure 15

Neutral adjective: Ordinary

List of queries:

- Query positive adjective + subject: Pure + Beach (labeled positive), see Figure 16
- Query negative adjective + subject: Polluted + Beach (labeled negative), see Figure 17
- Query neutral adjective + subject: Ordinary + Beach (labeled neutral), see Figure 18



Figure 11. Images retrieved from the query: Ferocious + Animal



Figure 13. Images retrieved from the query: Mildness

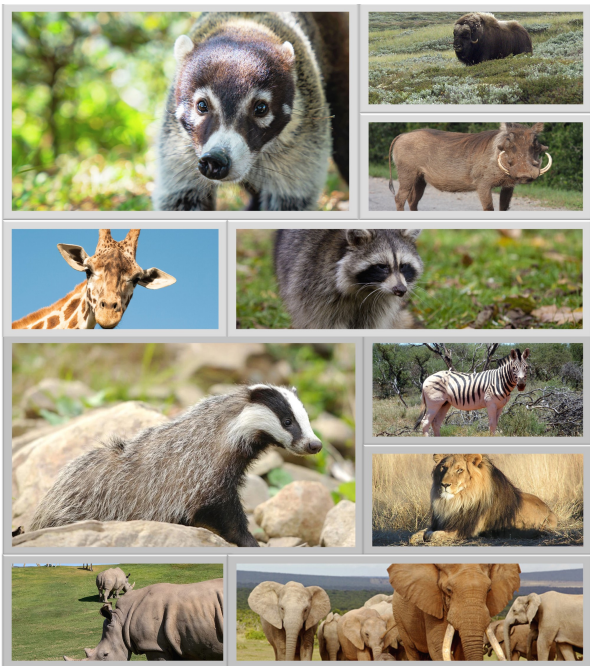


Figure 12. Images retrieved from the query: Neutral + Animal



Figure 14. Images retrieved from the query: Ferocity

- Query for positive artistic images: Purity (labeled positive), see Figure 19
- Query for negative artistic images: Pollution (labeled negative), see Figure 20

4. Ethical and Copyright Considerations for Human Subjects in the Dataset

In this section, we evaluate the usability of the images in the dataset, particularly those depicting individuals, from both ethical and copyright perspectives. All images in our

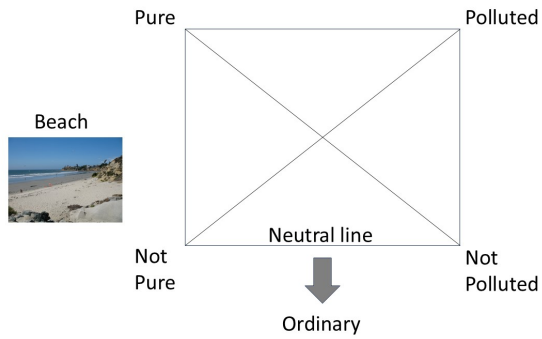


Figure 15. Example III: semiotic square application



Figure 16. Images retrieved from the query: Pure + Beach

dataset were carefully sourced from publicly available websites, ensuring that their licenses or usage conditions permitted free use and sharing for research purposes. Special attention was given to images of individuals, ensuring that the licenses under which they were distributed permitted their use for research purposes in compliance with ethical standards. The images were obtained through three sources: the Google Images search engine, Unsplash.com, and Pexels.com. For images downloaded from Google Images, we strictly adhered to licenses permitting free use for research purposes, including those depicting individuals. To guarantee compliance, we filtered images based on license type using Google’s tools and verified that the licenses indeed allowed usage for research purposes. Further details on how



Figure 17. Images retrieved from the query: Polluted + Beach

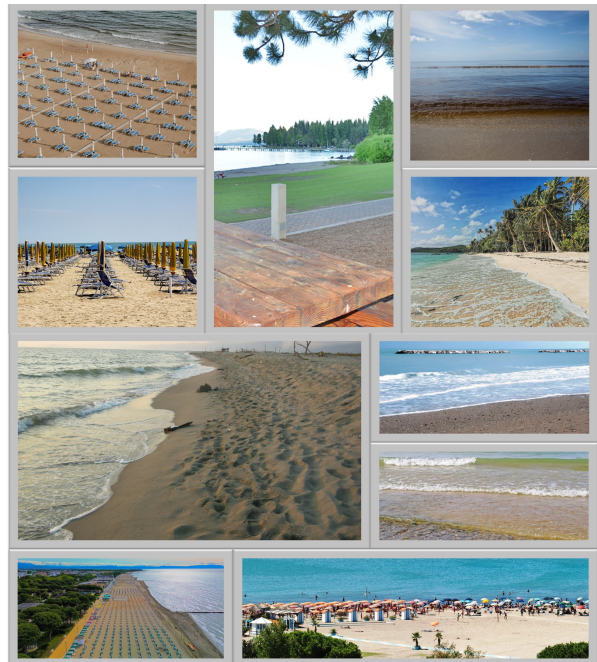


Figure 18. Images retrieved from the query: Ordinary + Beach

these licenses are defined and enforced can be found on Google Images’ copyright and usage rights page [2]. Images without clear licensing information or with restrictions on use and sharing for research purposes were excluded from our dataset.

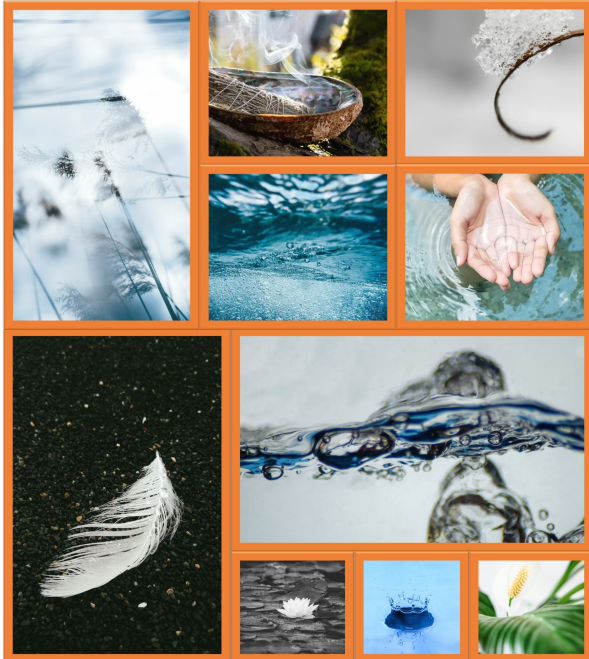


Figure 19. Images retrieved from the query: Purity



Figure 20. Images retrieved from the query: Pollution

For images obtained from Pexels, the platform explicitly states that the images they host are freely usable and shareable for research purposes, including those depicting individuals [1]. Similarly, images from Unsplash.com were selected only if their licenses allowed free use and sharing for research purposes. Images that had restrictions on

use or sharing were not downloaded and are not part of our dataset. [3].

The dataset was created exclusively for research purposes, ensuring that all images included comply with ethical standards and copyright regulations for such use.

5. Dataset download

In this section, we provide the link to download our dataset, the model trained with the CLIP Image Encoder architecture on our dataset, and the code to run it. The link is provided anonymously to comply with the anonymity rules for articles.

Link: https://mega.nz/folder/ncVxETpI#j1NMDDSGK_chAAmnJt3LbQ

The dataset will be made freely accessible for research purposes upon the publication of the article.

References

- [1] Pexels.com. <https://pexels.com/>. Accessed: 29/7/2023 - 30/3/2024.
- [2] Search results on Google. <https://support.google.com>. Accessed: 29/7/2023 - 30/3/2024.
- [3] Unsplash.com. <https://unsplash.com/>. Accessed: 29/7/2023 - 30/3/2024.
- [4] Steven Bird, Ewan Klein, and Edward Loper. *Natural Language Processing with Python*. O'Reilly Media Inc., 2009.
- [5] Andrea Esuli and Fabrizio Sebastiani. Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC 2006)*, pages 417–422, Genoa, Italy, 2006. European Language Resources Association (ELRA).
- [6] Merriam-Webster. Merriam-webster thesaurus. <https://www.merriam-webster.com/thesaurus>, 2024. Accessed: 29/7/2023 - 30/3/2024.