Automatically generating context-based alternative text using artificial intelligence techniques

1. United States Patent:12182525
2. Date of Patent: December 31, 2024
3. Inventor(s)
    a. **Beshara; Shary (Cairo, EG)**
    b. **Seibel; Amy (Newton, MA)**
    c. **Uppala; Kranti (Milford, MA)**
    d. **Bruno; Eric (Shirley, NY)**
4. Abstract: Methods, apparatus, and processor-readable storage media for automatically generating context-based alternative text using artificial intelligence techniques are provided herein. An example computer-implemented method includes generating text captions for an image derived from a web page by processing the image using an artificial intelligence-based image captioning model; determining context information pertaining to the image by processing the image using an artificial intelligence-based context and emotion recognition library; generating context-based alternative text for at least a portion of the image by processing, using at least one artificial intelligence-based alternative text generation model, at least a portion of one or more of the generated text caption(s) for the image and the determined context information pertaining to at least a portion of the image; and performing one or more automated actions based on the generated context-based alternative text.
5. BACKGROUND
    a. A significant number of websites are inaccessible to blind and visually-impaired users. For example, sighted people are often more attracted to images than words, which leads website designers to rely on images that, in many cases, blind and visually-impaired users cannot effectively interpret. Alternative text, also commonly referred to as ALT text, refers to a description of an image that can be written, for example, using hypertext markup language (HTML) elements. The use of ALT text attempts to render websites more accessible for users. However, in conventional ALT text approaches, human users typically must determine and manually enter the ALT text for a given image, which often results in errors and delays.

6. SUMMARY
    a. Illustrative embodiments of the disclosure provide techniques for automatically generating context-based alternative text using artificial intelligence techniques. An exemplary computer-implemented method includes generating one or more text captions for an image derived from a web page by processing at least a portion of the image using at least one artificial intelligence-based image captioning model, and determining context information pertaining to at least a portion of the image by processing one or more portions of the image using at least one artificial intelligence-based context and emotion recognition library. Additionally, the method includes generating context-based alternative text for at least a portion of the image by processing, using at least one artificial intelligence-based alternative text generation model, at least a portion of one or more of the one or more generated

text captions for the image and the determined context information pertaining to at least a portion of the image, and performing one or more automated actions based at least in part on the generated context-based alternative text.

b. Illustrative embodiments can provide significant advantages relative to conventional alternative text approaches. For example, problems associated with errors and delays are overcome in one or more embodiments through automatically generating context-based alternative text using artificial intelligence techniques.

c. These and other illustrative embodiments described herein include, without limitation, methods, apparatus, systems, and computer program products comprising processor-readable storage media.