June 2021

AI-based Image Synthesis for Enriched Search and Shopping

Emmanouil Koukoumidis

Shiblee Hasan

Joseph Edwin Johnson Jr.

AI-based Image Synthesis for Enriched Search and Shopping

ABSTRACT

Across numerous applications, notably in search and shopping for unique items, humans are constrained by what has already been built or designed. This disclosure describes techniques that leverage natural language-based, deep-learning image synthesis to deliver enhanced product search via services such as search engines or e-commerce websites. The synthetically generated products can be custom manufactured upon order. Unconstrained by real world objects, the techniques deliver to the search engine or e-commerce user synthetic objects based on text descriptions provided by the user.

KEYWORDS

- Image synthesis
- Video synthesis
- AI-generated object
- Machine generated image

- Deep learning
- Custom manufacturing
- Three-dimensional printing
- 3D printing

BACKGROUND

Across numerous applications, notably in search and shopping for unique items, humans are constrained by what has already been built or designed. For example, if a user searches or shops for "an armchair in the shape of an avocado" then the search results are likely to be very limited. A similar search for a "table that looks like a fish" returns very few relevant results as well. If the search is even more fine-grained e.g. "table that looks like a red fish" then hardly any relevant results are returned. This significantly constrains the ability of users to find unique items that fit their needs, let alone to fully conceptualize and visualize how such items would look. It is

useful for an online shopper to visualize such items before proceeding to order them from

retailers that can source or custom-manufacture them.
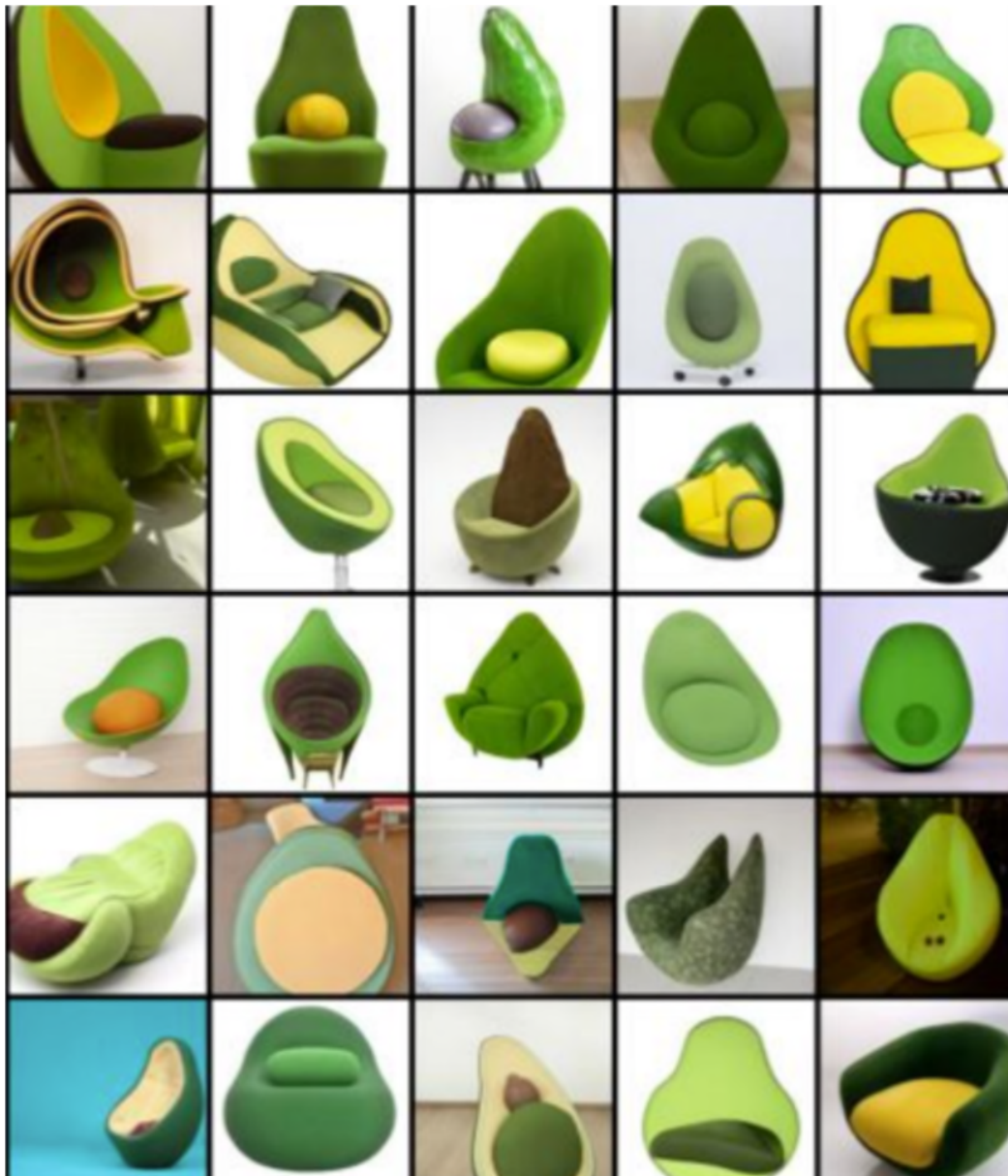
Fig. 1: Designs for an imaginary object synthesized by a deep learning model [1] for the natural-language prompt "an armchair in the shape of an avocado"

Fig. 1 illustrates designs for an imaginary object synthesized by a deep learning model

[1] for the natural-language prompt "an armchair in the shape of an avocado." Such technology

can effectively combine two unrelated concepts (armchair, avocado) based on directions provided in the prompt ("in the shape of").

DESCRIPTION
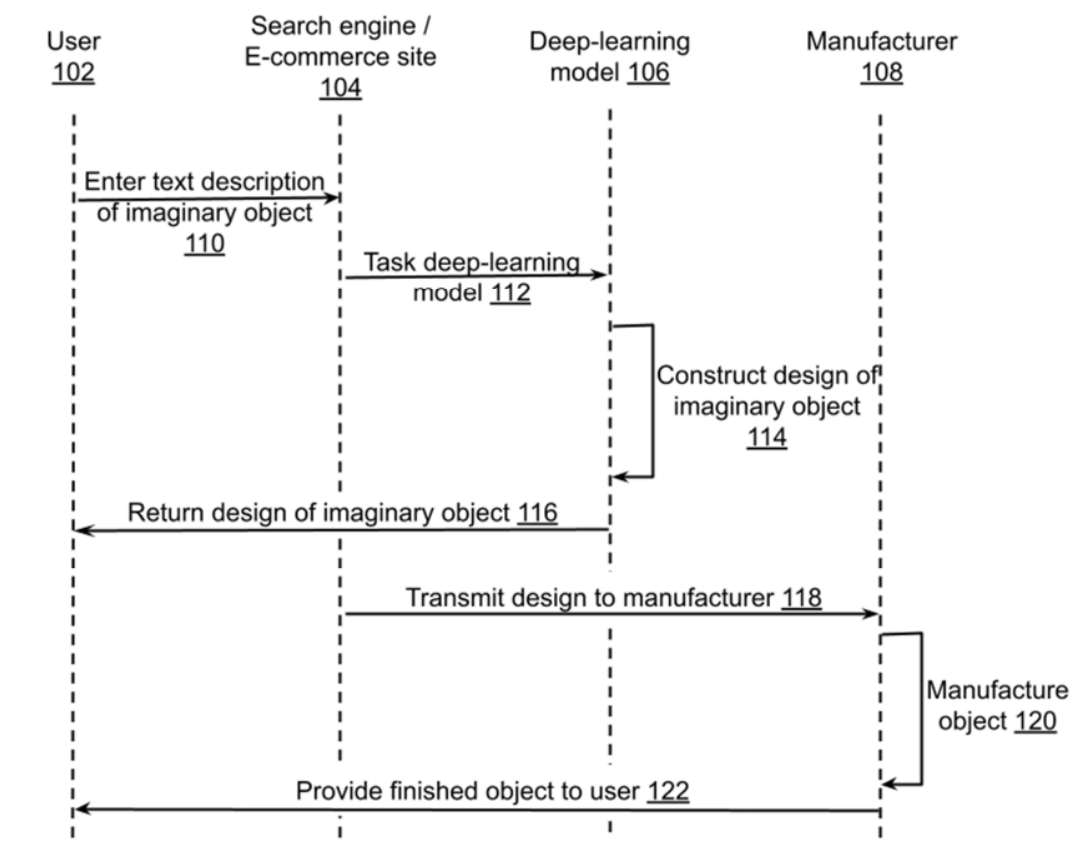
Fig. 1: Artificial Intelligence (AI) based Image Synthesis for Enriched Search and Shopping

Fig. 1 illustrates artificial intelligence (AI) based synthesis for enriched search and shopping, per the techniques of this disclosure. A user (102) provides a search engine, an e-commerce site or other service (104) a text description of a presently imaginary object (110), e.g., "a penguin doll shaped like a spacecraft," "a nightstand themed as a tree," etc.

After determining that the search phrase doesn't correspond to any extant object or image, the service utilizes a deep-learning model (106) to construct the imaginary object (112). The deep-learning model synthesizes from the text description an image or design of the

imaginary object (114). In the case of a search query, the design or image of the imaginary object

is returned to the user (116), alongside images of real objects that best match the query. In the

case of a shopping query followed by an order, the design can be transmitted to a manufacturer

(118), which can utilize a three-dimensional (3D) printer or other suitable technology to

manufacture the object. The manufacturer (108) manufactures the object (120). The finished

object is returned to the user (122). In the cases where a synthetic object is returned by the

service, the results provided to the user are clearly marked as being generated by AI.

Fig. 2: An example e-commerce site that incorporates AI-generated objects for custom manufacture and sale

Fig. 2 illustrates an example e-commerce site that incorporates AI-generated objects for custom manufacture and sale. A user enters a search phrase for an as-yet uncreated, imaginary product (202), e.g., a honeycomb-patterned phone case. The user is offered the option of including in the search or product results AI-generated objects (208). Among the search results are real, existing products (206) alongside as-yet nonexistent, AI-generated products (204). The AI-generated products are marked as such and are available for custom manufacture through preorder.

Some additional example use cases include:

- *Synthesized videos*: Since videos are essentially rapid sequences of still images, the described techniques can be utilized to serve video searches for synthetic scenes. For synthetic video creation, a frame can be conditioned on and contiguous with its previous frames. A user can search for a hitherto uncreated video, e.g., "a winged, fluorescent, parachute-like creature wades its way through a gaseous, high-viscosity exo-planet and chances upon a human object sent there in 1976," or "a 360-degree review of an avocado-shaped armchair." The search results can include synthetic, generated videos that match the text description, and they can also include real videos matching the text description as closely as possible. For example, a text search of "dodo flying" can return synthetic videos of the flightless, extinct bird flying, as well as nearly flightless, extant birds, e.g., turkeys, actually flying. Synthetic images and videos can be watermarked in such a way that other individuals/machines can identify them as being of synthetic origin.

- *Combining trends*: Trends from social media can be automatically combined by the deep-learning model to present to users synthetic objects as featured products. For example, if trending topics on social media are "glass terrarium" and "coffee table" then a synthetic

object "coffee table with integrated terrarium" can be auto-generated by a deep-learning model and presented to the user as featured products. This can pique user interest, and they can preorder such products. Conversely, a synthetic object may itself set off a trend; this can alert manufacturers to its popularity, whereupon they can prime their factories to fill the demand.

- *Creative professions*: Creative professionals, e.g., graphic artists, fashion designers, architects, cartoonists, interior designers, filmmakers, sculptors, social-media influencers, etc., can search for objects of their imagination and have an e-commerce site return them via a custom manufacturer. An interior designer can create, e.g., a rainforest-themed bedroom, where the furniture is shaped and colored like a rainforest. A filmmaker can create, e.g., a set for rodent-aliens on an exoplanet. A sculptor can create, e.g., abstract forms that lend themselves to variegated meaning. A fashion designer can experiment, e.g., with bell-bottoms patterned with avocados. A social-media influencer can create a video by simply describing it verbally.

- *Auto-completion of search images, or improving existing searches:* The synthetic object and video search results generated by the described techniques can be used as a seed to auto-complete search results or improve existing searches, as follows. Prior to completion of the search query, or just after, one or more synthetic images or videos can be presented to the user for them to choose from as the one that matches closest to their search intent. Such an image or video can be used to refine the search results using image and/or text caption features (embeddings) extracted from the user-selected image. Even without the user-selection step, the synthetic image may be useful for comparing/refining search results based on the existing query.

- *Auto-correction of user queries:* If the result of a user search has insufficient relevance or diversity, synthetic search results can be generated to improve the relevance, diversity and thus the richness of the overall user experience. For example, if a user types, "avocado armchair," and no satisfying pictures are found, a picture of one or more synthesized avocado armchairs can be displayed alongside a note that the displayed images are synthetic and do not correspond to real objects.

In this manner, the techniques of this disclosure leverage natural language-based, deep-learning image synthesis to deliver an enhanced product search for search engines (web search, online shopping search) and custom manufacturing. Unconstrained by real-world objects, the techniques deliver to the search-engine or e-commerce user synthetic objects based on text descriptions provided by the user. Search trends for imaginary objects or videos can be presented to advertising/search/retail partners to augment or improve items they offer for sale, services, etc. The techniques can unlock substantial market opportunities for search engines that focus on web/image/product search; for video-sharing and photo-sharing sites; for online shops and e-commerce sites that want to expand their product portfolio; for creative professionals, e.g., graphic artists, interior designers, fashion designers, architects, cartoonists, sculptors, etc.; for manufacturers of custom products; etc.

CONCLUSION

This disclosure describes techniques that leverage natural language-based, deep-learning image synthesis to deliver enhanced product search via services such as search engines or e-commerce websites. The synthetically generated products can be custom manufactured upon order. Unconstrained by real world objects, the techniques deliver to the search engine or e-commerce user synthetic objects based on text descriptions provided by the user.

REFERENCES

[1] "DALL·E: Creating Images from Text" https://openai.com/blog/dall-e/ accessed May 29, 2021.

[2] https://liamp.substack.com/p/my-gpt-3-blog-got-26-thousand-visitors accessed May 29, 2021.