PECKHAMIA 288.1, 5 January 2023, 1—30
LSID urn:lsid:zoobank.org:pub:B409A798-88CC-4918-9AF5-2F31D2065DF6 (registered 30 DEC 2022)
ISSN 2161—8526 (print)
ISSN 1944—8120 (online)

Archetypes of the *jumping spider* (Araneae: Salticidae) as derived by intelligent machines [1]

David E. Hill [2]

[1] Definitions for key terms employed in this paper are given in Appendix 1.
[2] 213 Wild Horse Creek Drive, Simpsonville, South Carolina 29680, USA, *email* platycryptus@yahoo.com

Recent papers have advanced the hypothesis that mimicry of salticid spiders by a variety of insects can provide protection to those insects against the attack of salticid spiders (Hill, Abhijith & Burini 2019; Hill 2022). In effect, as predator on an insect population, salticid spiders selectively prey on individuals less able to represent themselves as salticid spiders, based on an *archetypal image* or *engram* held by those spiders (Figure 1.1-1.2).

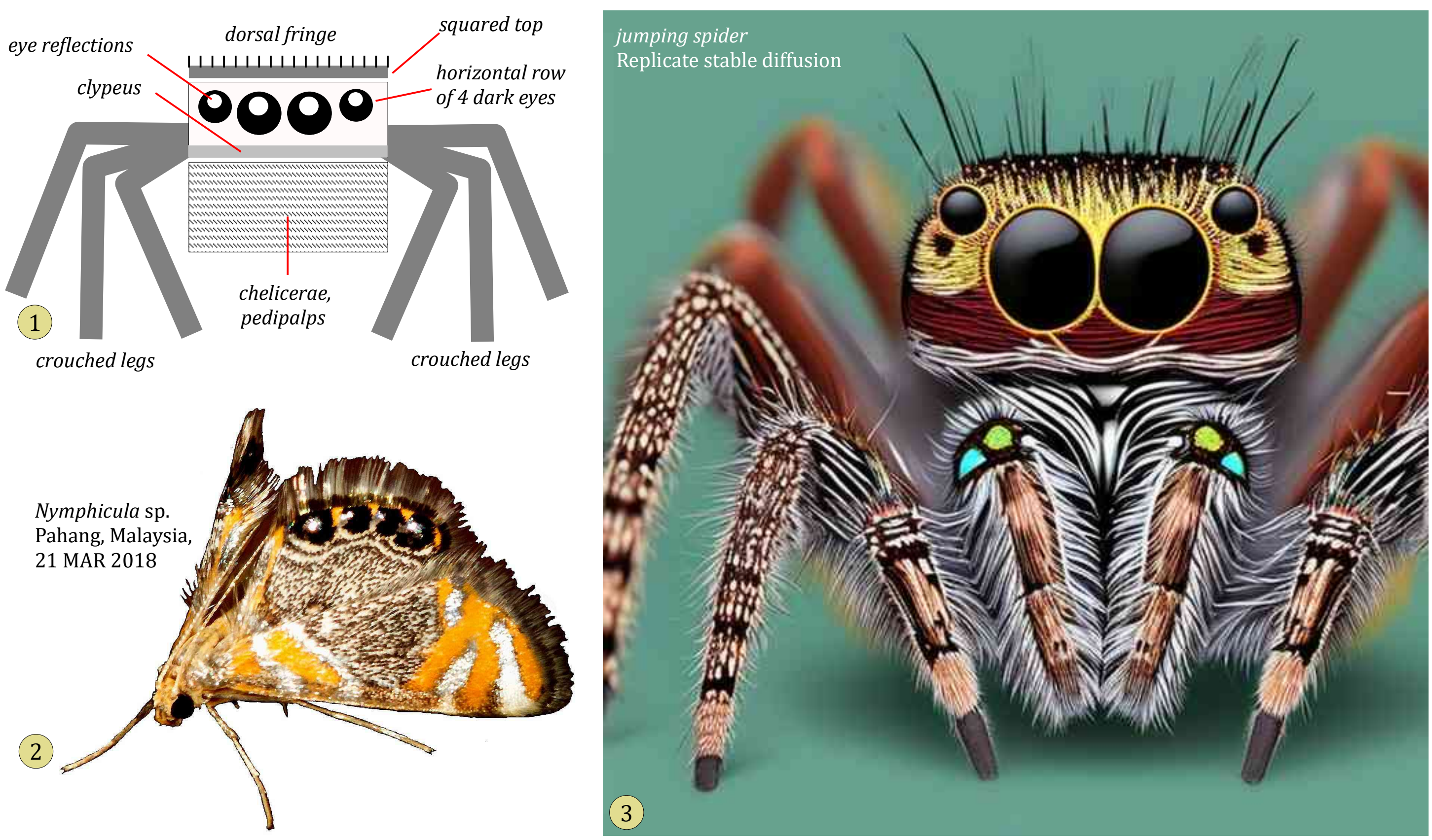

Figure 1. 1, Hypothetical archetype of a salticid spider, showing key features that may be used by another salticid to recognize this image as a salticid (after Hill 2022). **2,** Example of a crambid moth that displays the key features of this archetype on its wings. **3,** Image generated in response to the text input *jumping spider, highly detailed* by an intelligent machine (Replicate stable diffusion; see Table 1, #7). Can intelligent machines provide us with insight into the nature of the salticid archetype?

Here I will consider the possibility that a new generation of intelligent machines capable of generating images based on a text input (*text to image*), recently made accessible to the public, can help us to understand the nature of the salticid archetype of salticid spiders. For purposes of this study, a series of images were obtained by presenting a simple text message (usually *jumping spider, highly detailed*) multiple times to each of a series of computer systems (*engines*) that have the ability to convert text statements to images (Table 1). As can be seen from the results (Figures 1.3, 2-30), a diverse set of images was produced as a result.

Table 1. Online engines used to generate images of jumping spiders from text.

	engine	web site	technology	technology references	figures
1	Pixray vqgan	https://replicate.com/pixray/text2image	CLIP, VQGAN	Radford et al. 2021; Esser, Romback & Ommer 2021	2
2	DeepAI	https://deepai.org/machine-learning-model/text2img			3
3	NightCafe Stable Diffusion, photo	https://creator.nightcafe.studio/	stable diffusion	Rombach et al. 2022	4-5
4	Stable Diffusion Playground	https://stablediffusionweb.com/#demo	stable diffusion		6-8
5	Canva	https://www.canva.com/			9-10
6	Huggingface Stable Diffusion 2.1	https://huggingface.co/spaces/stabilityai/stable-diffusion	stable diffusion		11-14, 24-30
7	Replicate Stable Diffusion	https://replicate.com/stability-ai/stable-diffusion	stable diffusion		1.3, 15-16
8	Starryai Argo 2	https://starryai.com/	stable diffusion		17-18
9	Simplified	https://simplified.com/ai-image-generator/	CLIP, diffusion		19-20
10	OpenAI DALL-E 2	https://openai.com/dall-e-2/	CLIP, diffusion, GLIDE	Nichol et al. 2022	21-23
11	Deep Dream Generator	https://deepdreamgenerator.com/	stable diffusion		31-32

When processing, the Pixray vggan engine (#1) allowed the user to view successive iterations as each image was drawn. However, unlike the other engines reviewed here, the images that were produced were quite primitive and depicted little variation in form (Figure 2). Nonetheless they still included many key features of the salticid archetype, to include a horizontal row of eyes, a dorsal fringe, a clypeus, and several legs.

Figure 2. Responses to the *jumping spider, high detail* text prompt by the Pixray vqgan engine (Table 1: #1). Although primitive, elements of the salticid archetype shown in Figure 1.1, including a horizontal row of eyes, a clypeus, and a fringe above the eyes, can be seen here.

The DeepAI (#2, Figure 3) and NightCafe stable diffusion (#3, Figures 4-5) engines produced a greater variety of images that were much more like salticid spiders, but with significantly less accuracy than the subsequent engines listed in Table 1.

Figure 3. Responses to the *jumping spider, high detail* text prompt by the DeepAI engine (#2).

Figure 4. Responses to the *jumping spider, high detail* text prompt by the NightCafe Stable Diffusion engine (#3).

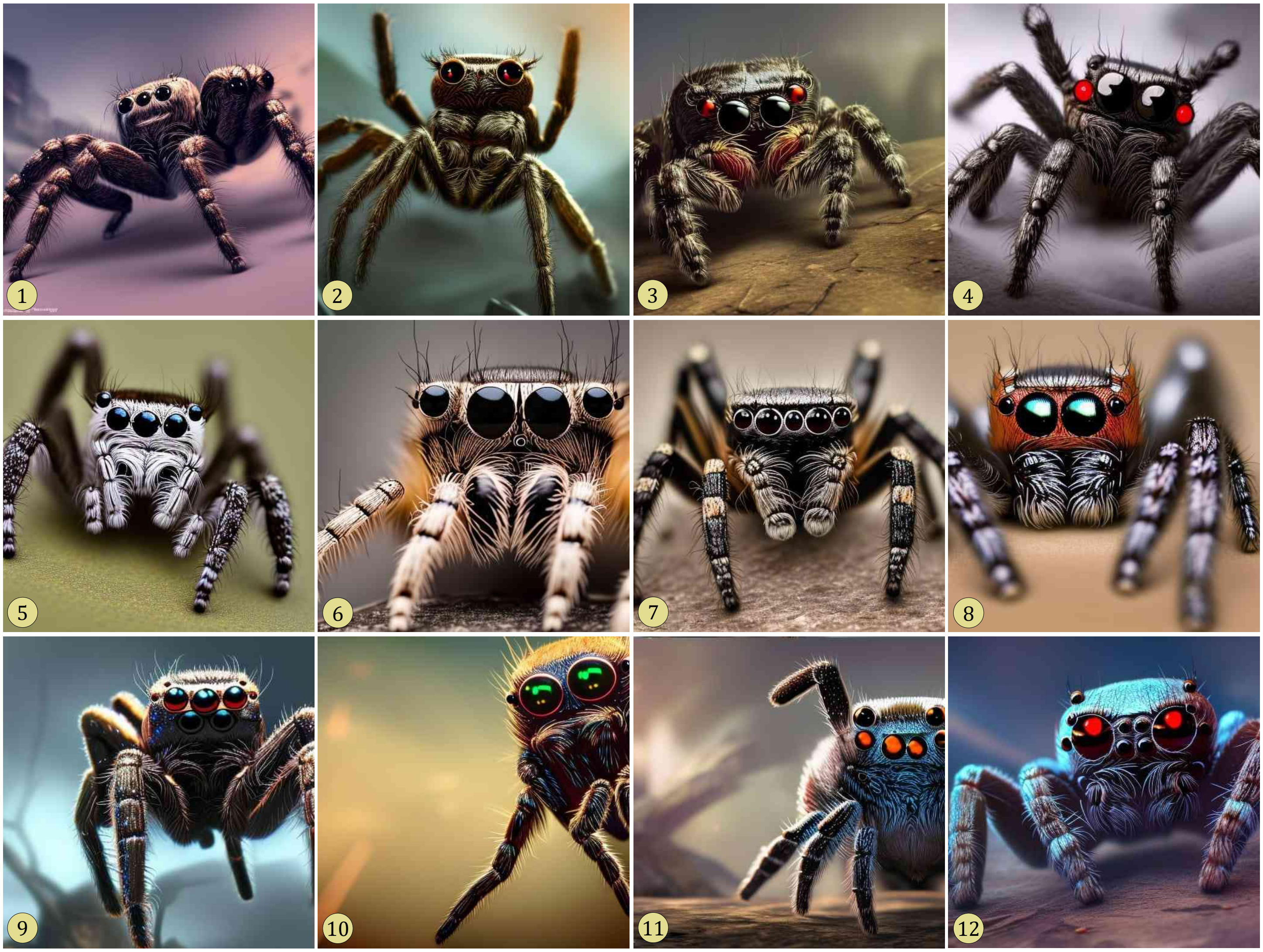

Figure 5. Responses to the *jumping spider, high detail* text prompt by the NightCafe Stable Diffusion engine (#3).

A series of intermediate engines (Stable Diffusion Playground, #4, Figures 6-8; Canva, #5, Figures 9-10; Huggingface stable diffusion, #6, Figures 11-14; Replicate Stable Diffusion, #7, Figures 1.3, 15-16) and Starryai Argo 2, #8, Figures 17-18) tended to produce much better images of salticid spiders with many variations. The Simplified (#9, Figures 19-20) and OpenAI DALL-E 2 (#10, Figures 21-22) consistently produced remarkably accurate images, but with far less variation.

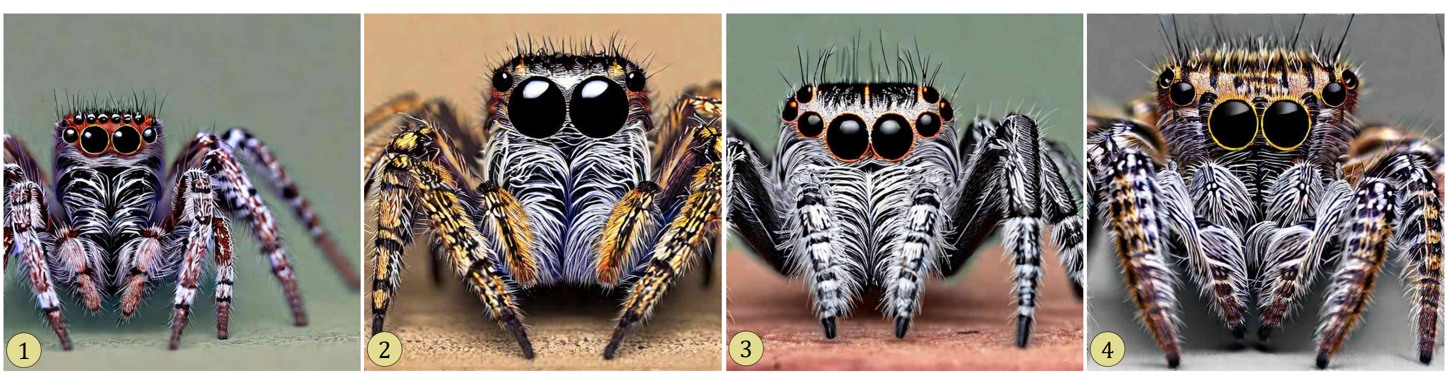

Figure 6. Responses to the *jumping spider, high detail* text prompt by the Stable Diffusion Playground engine (#4).

Figure 7. Responses to the *jumping spider, high detail* text prompt by the Stable Diffusion Playground engine (#4).

Figure 8. Responses to the *jumping spider, high detail* text prompt by the Stable Diffusion Playground engine (#4).

Figure 9. Responses to the *jumping spider, high detail* text prompt by the Canva engine (#5).

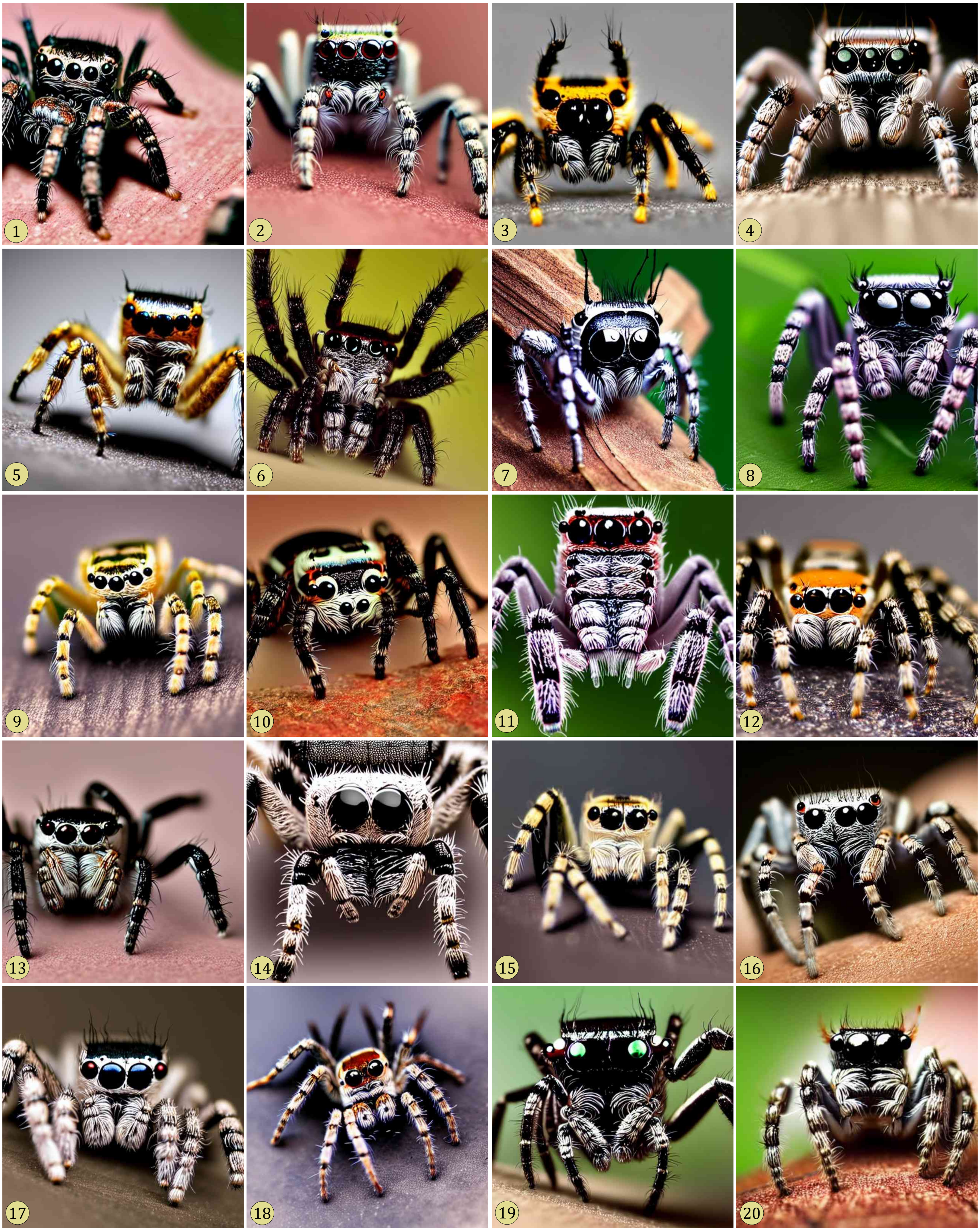

Figure 10. Responses to the *jumping spider, high detail* text prompt by the Canva engine (#5).

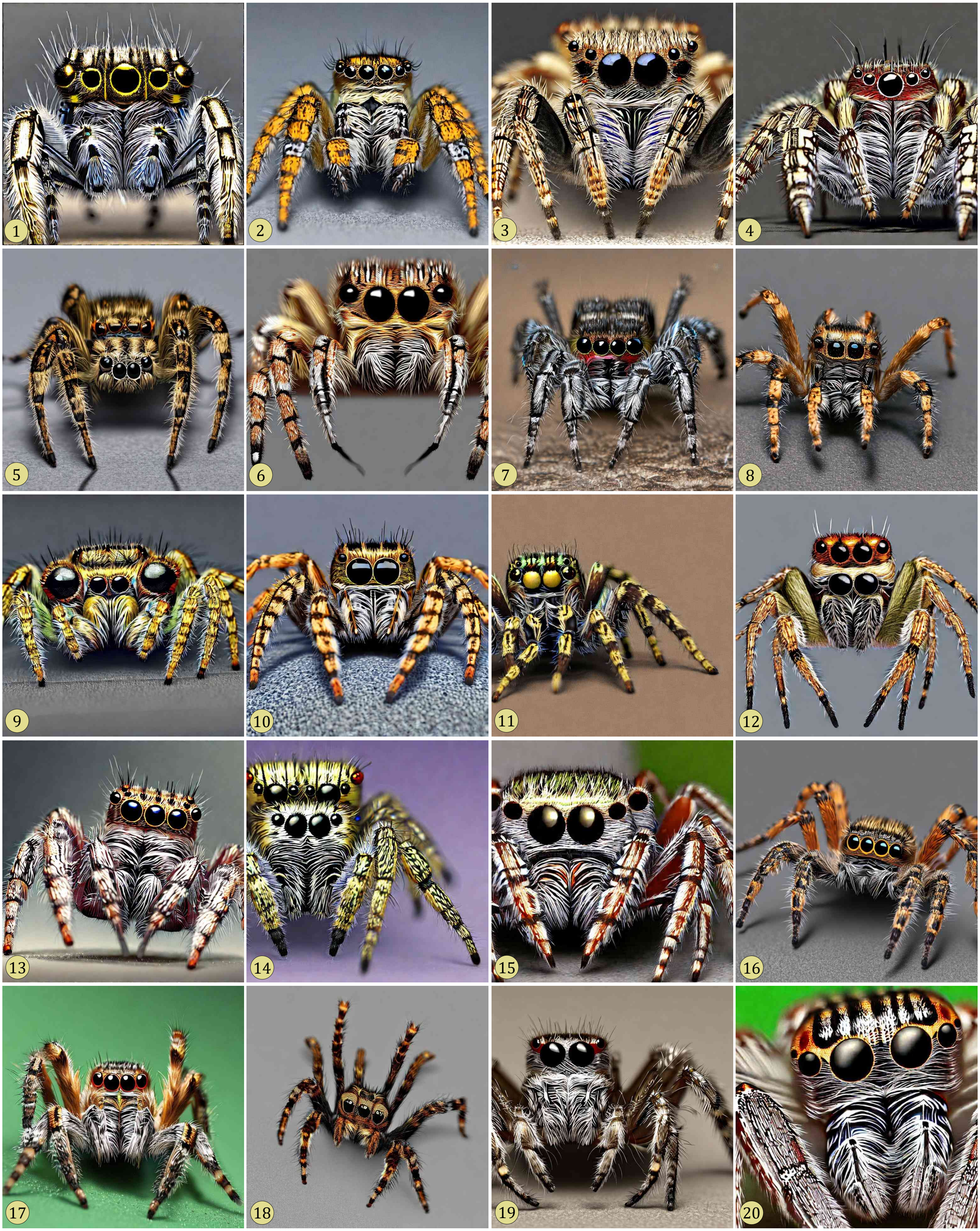

Figure 11. Responses to the *jumping spider, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

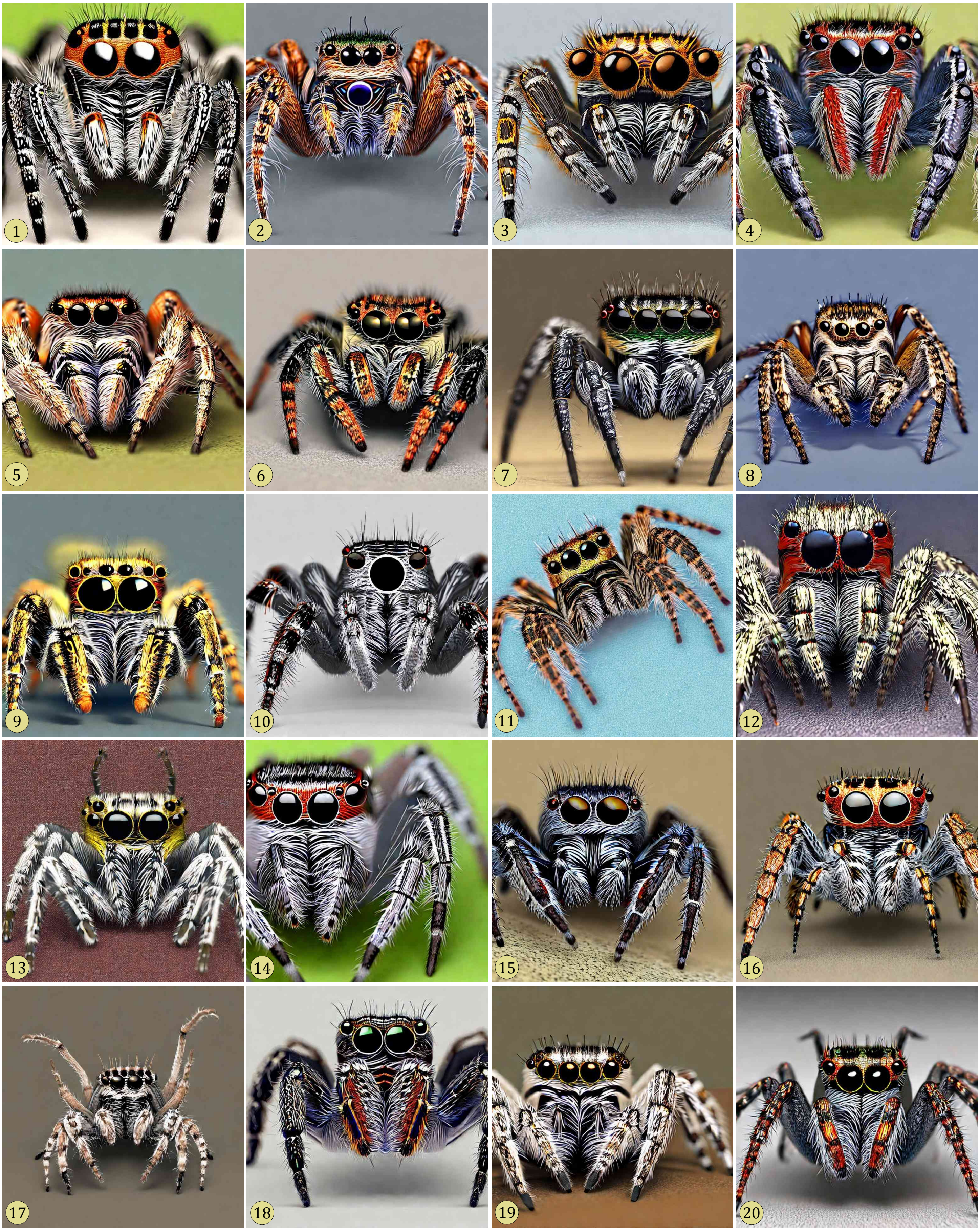

Figure 12. Responses to the *jumping spider, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 13. Responses to the *jumping spider, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 14. Responses to the *jumping spider, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 15. Responses to the *jumping spider, high detail* text prompt by the Replicate Stable Diffusion engine (#7).

Figure 16. Responses to the *jumping spider, high detail* text prompt by the Replicate Stable Diffusion engine (#7).

Figure 17. Responses to the *jumping spider, high detail* text prompt by the Starryai Argo 2 engine (#8).

Figure 18. Responses to the *jumping spider, high detail* text prompt by the Starryai Argo 2 engine (#8).

Figure 19. Responses to the *jumping spider, high detail* text prompt by the Simplified engine (#9).

Figure 20. Responses to the *jumping spider, high detail* text prompt by the Simplified engine (#9).

Figure 21. Responses to the *jumping spider, high detail* text prompt by the OpenAI DALL-E 2 engine (#10).

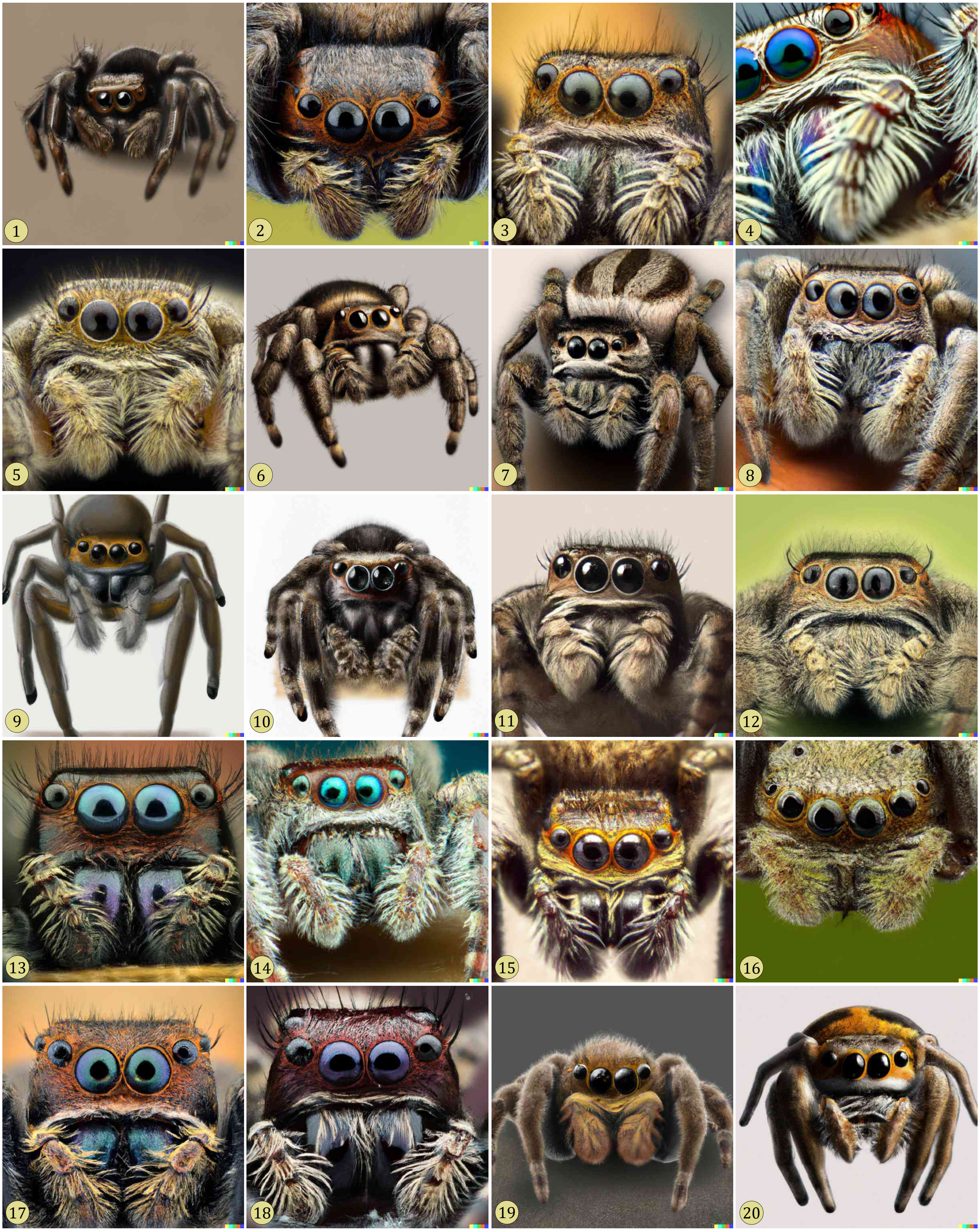

Figure 22. Responses to the *jumping spider, high detail* text prompt by the OpenAI DALL-E 2 engine (#10).

Figure 23. Responses to the *jumping spider, high detail* text prompt by the OpenAI DALL-E 2 engine (#10).

To further study the role of natural language in eliciting these responses, I also tried several other prompts using the Huggingface Stable Diffusion 2.1 engine (#6). In response to a *peacock spider, high detail* prompt, generated images resembled spiders with peacock-like colors and displays (Figure 24). In contrast, a *peacock jumping spider, high detail* prompt produced mostly good images of jumping spiders with bright, peacock-like colors (Figure 25). A *Phidippus*, *high detail* prompt produced a variety of spider-like images, mostly with at least one pair of large, forward facing eyes (Figure 26), but a *Phidippus jumping spider, high detail* prompt produced much more accurate images of salticids (Figure 27), similar to those produced in response to the basic *jumping spider, high detail* prompt (Figures 11-14). The *salticid, high detail* prompt produced a series of spider-like images (Figure 28), however not much like salticids. The *Salticidae, high detail* prompt (Figure 29) was much better at this. Finally a *Springspinne, high detail* prompt produced a series of arthropod-like images, but none that could be recognized as any kind of spider (Figure 30).

These results indicate that, whereas the relationship between *jumping spider* and images of salticids is well-trained into this system, some text may be much less effective or predictable. *Salticidae* is better-trained than is *salticid*. Curiously, whereas *Phidippus* tends to produce images suggestive of a jumping spider, the German term *Springspinne* is much less effective, although both produce arthropod-like images in any case. These differences simply reflect training of the system, not some fundamental limitation.

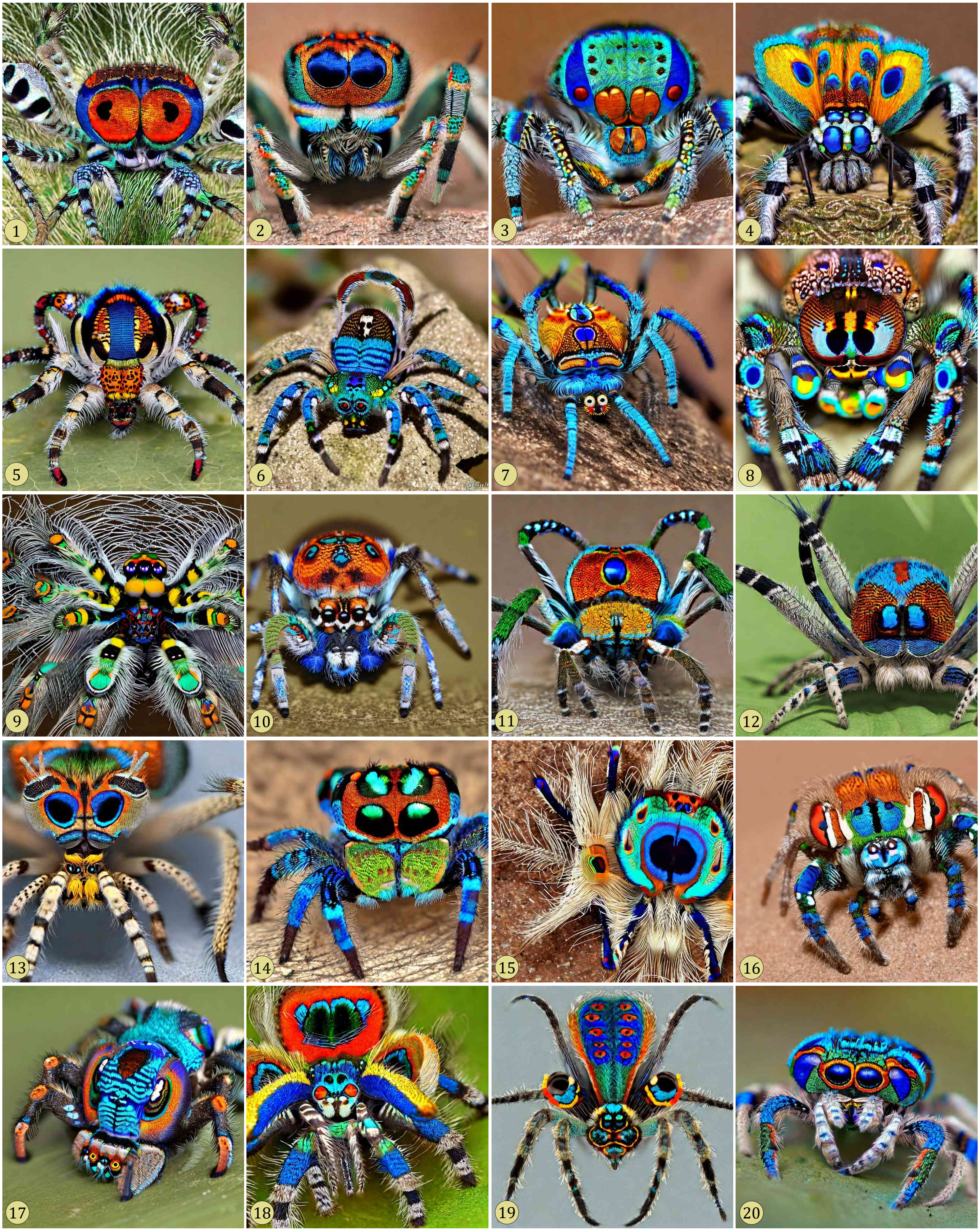

Figure 24. Responses to the *peacock spider, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 25. Responses to the *peacock jumping spider, high detail* prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 26. Responses to the *Phidippus, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 27. Responses to the *Phidippus jumping spider, high detail* prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 28. Responses to the *salticid, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 29. Responses to the *Salticidae, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Figure 30. Responses to the *Springspinne, high detail* text prompt by the Huggingface Stable Diffusion 2.1 engine (#6).

Discussion. Clearly a number of different *text to image* systems are successful at producing a reasonable image of a salticid, often quite variable but in general bearing the features that have been associated with a salticid archetype. How should we interpret this finding? Although there have been some recent efforts to model the features of object recognition by primates with deep neural network computers (e.g., Rajalingham et al. 2018), we still need to know much more about the neurobiology of jumping spiders for approaches of this kind to be meaningful.

What we do know, however, is that without specific programming to this effect, neural network computers are capable of learning the association between natural language (in our case the English term *jumping spider*) and general features derived from the images represented by this language. Here these features appear to include those that we have associated with a salticid archetype (Figure 1.1). By itself, this supports the salticid archetype hypothesis to the extent that respective features of that archetype can be independently derived through the objective observation of, and feature extraction from, a series of salticid images by the intelligent machine.

A series of recent technological advances have made this kind of machine intelligence possible. Since 2014, Generative Adversarial Networks (GAN) have been used successfully in a number of intelligent machines, most notably demonstrated by the ability of *AlphaGo Zero* to quickly teach itself to defeat the best human *Go* players (Dong, Wu & Zhou 2017). However, for the purposes of text to image generation, *diffusion models* (e.g., Choi et al. 2022; Saharia et al. 2022) have recently been more successful. These have, in some cases, been enhanced or modified through the addition of GAN features (Wang et al. 2022), or even more efficient *pretrained encoders* (Rombach et al. 2022). Diffusion-based systems are transformers, selecting and recording image features by determining and remembering the steps required to convert an image (associated with text) to a pixel array of pure, random Gaussian noise. To produce a new image from text input, these systems begin with a pixel array of random noise, and then reverse the sequence of steps that they have learned, to recompose images with features like those that they were trained on. As a result they produce a different image each time, but one that usually retains the selected features that they have learned. Many examples have been included in this paper to demonstrate just how varied these productions can be.

A discussion of terms relevant to this discussion is in order (see Appendix 1). Here I do not use the popular term *artificial intelligence* with respect to information processing by machines (rather, *machine intelligence*), as the term *intelligence* is equally applicable to both organismal and machine behavior. Likewise, both machines and organisms may have the ability to acquire useful information (with both *learning* and *memory*). Notions used to describe objects that can be associated with our own natural language (e.g., *concept*, *semantic object, symbol*) are equally applicable to machines and organisms. The internal representations of concepts held by salticids (*archetypes* or *engrams*) are still quite mysterious to us, and in some cases, as with the salticid *archetype*, they appear to be *innate* (not acquired through learning). Important engrams related to conspecific or mate recognition, or prey recognition (Edwards & Jackson 1993; Bartos 2022) by the jumping spider may also be largely innate.

Surrealism. Text to speech systems have recently spurred the production of a vast amount of art that can be termed surrealist (unreal, but perhaps reflecting an underlying reality). As shown in Figures 31-32, the text input *jumping spider* can be combined with other terms (in this case *walking on vegetation*) as well as terms that describe an artistic style. These systems provide one more example of feature extraction and representation by machine intelligence, but their output may also have something to tell us about our own fascination with these creatures. Whether these images can also produce a supernormal response by jumping spiders remains to be determined.

Figure 31. Responses to the *jumping spider walking on vegetation* text prompt by the Deep Dream Generator engine (#11). One limitation of this and similar diffusion engines lies in their inability to consistently count serial objects (e.g., eyes in a row), but they are very good at creatively combining different text inputs to produce surrealistic images.

Figure 32. Responses to the *jumping spider walking on vegetation* text prompt by the Deep Dream Generator engine (#11).

Acknowledgements

I am most grateful to each group of developers that has given me access to their intelligent machines for purposes of this study, as listed in Table 1. I also thank Barbara Bauer, formerly a Director at U S West Advanced Technologies in Englewood, Colorado, for giving me responsibilities under our corporate sponsorship of early AI research at Carnegie-Mellon University in Pittsburgh, Pennsylvania, in the late 1980's. Discussions with the research team at Carnegie-Mellon ranged freely from neural networks to salticid behavior. I also thank Arnold Wijker for permitting the use of one of his photographs. All other images that appear here may be used under the terms of a CC BY 4.0 license, and in some cases their original versions may also be available online in the public domain.

References

Ai 2021. Shireen Ai. 13 JAN 2021. Dada and Surrealism, in brief. *Online at* https://www.researchgate.net/publication/348437078_ Dada_and_Surrealism_in_Brief

Bang & Shim 2018. Duhyeon Bang and Hyunjung Shim. Improved training of Generative Adversarial Networks using representative features. Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden PMLR 80, 2018.

Bartos 2022. Maciej Bartos. Visual prey categorization by a generalist jumping spider. The European Zoological Journal 89 (1): 1312-1324.

Chen et al. 2021. Hanting Chen, Yunhe Wang, Tianyu Guo, Chang Xu, Yiping Deng, Zhenhua Liu, Siwei Ma, Chunjing Xu, Chao Xu and Wen Gao. 8 NOV 2021. Pre-trained image processing transformer. Cornell University arXiv:2012.00364v4.

Choi et al. 2022. Jooyoung Choi, Jungbeom Lee, Chaehun Shin, Sungwon Kim, Hyunwoo Kim and Sungroh Yoon. 1 APR 2022. Perception prioritized training of diffusion models. Cornell University arXiv:2204.00227v1.

Dong, Wu & Zhou 2017. Xiao Dong, Jiasong Wu and Ling Zhou. 24 NOV 2017. Demystifying AlphaGo Zero as AlphaGo GAN. Cornell University arXiv:1711.09091v1.

Edwards & Jackson 1993. G. B. Edwards and Robert R. Jackson. Use of prey-specific predatory behaviour by North American jumping spiders (Araneae, Salticidae) of the genus *Phidippus*. Journal of Zoology, London 229: 709-716.

Esser, Romback & Ommer 2021. Patrick Esser, Robin Rombach and Björn Ommer. 23 JUN 2021. Taming transformers for high-resolution image synthesis, v. 3. Cornell University arXiv:2012.09841v3.

Goodfellow et al. 2014. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 10 JUN 2014. Generative Adversarial Networks. Cornell University arXiv:1406.2661v1.

Goodfellow et al. 2020. Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. NOV 2020. Generative Adversarial Networks. Communications of the ACM 63 (11): 139-144.

Gurney 1997. Kevin Gurney. An introduction to neural networks. UCL Press Limited, London.

Hill 2022. David E. Hill. 23 AUG 2022. A salticid archetype for salticid spiders. Peckhamia 275.1: 1-39.

Hill, Abhijith & Burini 2019. David E. Hill, Abhijith A. P. C. and João P. Burini. 6 FEB 2019. Do jumping spiders (Araneae: Salticidae) draw their own portraits? Peckhamia 179.1: 1-14.

Hovannisyan et al. 2021. Mariam Hovhannisyan, Alex Clarke, Benjamin R. Geib, Rosalie Cicchinelli, Zachary Monge, Tory Worth, Amanda Szymanski, Roberto Cabeza and Simon W. Davis. 19 JAN 2021. The visual and semantic features that predict object memory: Concept property norms for 1,000 object images. Memory & Cognition 49: 712-731.

Khan et al. 2022. Salman Khan, Muzammal Naseer, Munawar Hayat, Syed Waqas Zamir, Fahad Shahbaz Khan, and Mubarak Shah. 19 JAN 2022. Transformers in vision: a survey. Cornell University arXiv:2101.01169v5.

Liu et al. 2017. Weibo Liu, Zidong Wang, Xiaohui Liu, Nianyin Zeng, Yurong Liu and Fuad E. Alsaadi. 19 APR 2017. A survey of deep neural network architectures and their applications. Neurocomputing 234: 11-26.

Morizet 2020. Nicolas Moriset. Introduction to Generative Adversarial Networks. [Technical Report]. Advestis. 2020. hal-02899937.

Nichol et al. 2022. Alex Nichol, Prafulla Dhariwal, Aditya Ramesh, Pranav Shyam, Pamela Mishkin, Bob McGrew, Ilya Sutskever, and Mark Chen. 8 MAR 2022. Towards photorealistic image generation and editing with text-guided diffusion models, v.3. Cornell University arXiv:2112.10741v3.

Nielson 2019. Michael Nielson. Neural networks and deep learning, *online at:* http:// neuralnetworksanddeeplearning.com/

Parmar et al. 2018. Niki Parmar, Ashish Vaswani, Jakob Uszkoreit, Łukasz Kaiser, Noam Shazeer, Alexander Ku and Dustin Tran. 15 JUN 2018. Image transformer. Cornell University arXiv:1802.05751v3. Proceedings of the 35 th International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.

Radford et. al 2021. Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 26 FEB 2021. Learning transferable visual models from natural language supervision. Cornell University arXiv:2103.00020v1.

Rajalingham et. al. 2018. R. Rajalingham, E. B. Issa, P. Bashivan, K. Kar, K. Schmidt and J. J. DiCarlo. Large scale, high-resolution comparison of the core visual object recognition behavior of humans, monkeys, and state-of the art deep artifical neural networks. Journal of Neuroscience 38 (33): 7255-7269.

Rombach et al. 2022. Robin Rombach, Andreas Blattman, Dominik Lorenz, Patrick Esser and Björn Ommer. 13 APR 2022. High resolution image synthesis with latent diffusion models, v.2. Cornell University arXiv:2112.10752v2.

Saharia et al. 2022. Chitwan Saharia, William Chan, Saurabh Saxenay, Lala Liy, Jay Whangy, Emily Denton, Seyed Kamyar Seyed Ghasemipour, Burcu Karagol Ayan, S. Sara Mahdavi, Rapha Gontijo Lopes, Tim Salimans, Jonathan Hoy, David J Fleety, and Mohammad Norouzi. 3 MAY 2022. Photorealistic text-to-image diffusion models with deep language understanding. Cornell University arXiv:2205.11487v1.

Vidya 2018. T. N. C. Vidya. 10 SEP 2018. Supernormal stimuli and responses. Resonance 23: 853-860.

Zamorski et al. 2019. Maciej Zamorski, Adrian Zdobylak, Maciej Zięba and Jerzy Świątek. 16 MAR 2019. Generative adversarial networks: recent developments. Cornell University arXiv:1903.12266v1.

Wang et al. 2022. Zhendong Wang, Huangjie Zheng, Pengcheng He, Weizhu Chen, and Mingyuan Zhou. 11 OCT 2022. Diffusion-GAN: training GANs with diffusion. Cornell University arXiv:2206.02262v3.

Zeng & Kwong 2022. Chao Zeng and Sam Kwong. 23 MAR 2022. Learning transformer features for image quality assessment. Cornell University arXiv:2112.00485v2.

Appendix 1. Key terms used in this paper, with references

animal learning	acquisition of information by animals that can be used to modify behavior
animal memory	memory held by animals
archetypal image	image that can be recognized by the use of innate memory
artificial intelligence (AI)	older term for machine intelligence (MI), based on the view that machines are not really intelligent
CLIP	Contrastive Language-Image Pre-Training (Radford et al. 2021)
cognition	loaded term (multiple meanings) that should probably be avoided; some use this a a synonym of *recognition*, others to suggest *deep understanding*
concept	fundamental representation of any object or thing that can be handled by an information processing system
deep neural network	a complex neural network with multiple layers, usually with hidden layers (Liu et al. 2017; Rajalingham et al. 2018)
Diffusion-GAN	GAN trained with image diffusion (Wang et al. 2022)
diffusion model	model of visual object based on reversing steps required to turn it to noise (Choi et al. 2022; Saharia et al. 2022)
engram	an internal representation of an object that can be suggested by the behavior of an information system
GAN	Generative Adversarial Network (Goodfellow et al. 2014, 2020; Bang & Shim 2018; Zamorski et al. 2019; Morizet 2020)
GLIDE	Guided Language to Image Diffusion for Generation and Editing (Nichol et al. 2022)
information	facts, symbolic and other representations of fact
information processing	the collection, processing, storage, and utilization of information by a system
information system	living entity or machine that is capable of information processing
innate memory	information that is built into a system through development or manufacturing, but not acquired through a learning process
intelligence	ability to process information in a useful manner
intelligent animal	any animal with the ability to process information
intelligent machine	any machine with the ability to process information
latent diffusion model	use of pretrained autoencoders to support diffusion synthesis (Rombach et al. 2022)
learning	acquisition of information by an information system
machine intelligence (MI)	preferred term for what has been called artificial intelligence, based on the view that machines can be intelligent
machine learning	acquisition of information by machines that can be used to modify behavior
machine memory	memory held by machines
memory	retention of information that can be used to modify behavior at a later time
natural intelligence (NI)	intelligence of animals and other organisms that is not machine intelligence
natural language	a symbolic system acquired and used by humans to communicate
neural network computer	assembly of simple processing units that can learn through modification of the connections between units (Gurney 1997; Nielson 2019)
object	general term for anything that is subject to a symbolic representation
object recognition	ability to associate an object with its symbolic or internal representation (Hovhannisyan et al. 2021)
salticid archetype	archetypal image of a salticid spider held by another salticid spider (Hill 2022)
semantic model	set of representations for objects and their relationships that may be held in memory
semantic object	representation of any object in a semantic model
semantic relationship	relationship between semantic objects, for example, A *is a* B, or A *is a characteristic* of B
supernormal	a response that is stronger than the normal response that has evolved (Vidya 2018)
surrealism	in art, the exploration of the unreal thought to respresent an underlying reality (Ai 2021)
symbol	abstract representation of a semantic object, for example, a word or computer code
text to image generation	production of an image from text input
transformer	computer system produces an internal representation of an image (Parmar et al. 2018; Chen et al. 2021; Zeng & Kwong 2022; Khan et al. 2022)
VQGAN	Vector Quantized Generative Adversarial Networks (Esser, Romback & Ommer 2020)