A Misleading Gallery of Fluid Motion by Generative Artificial Intelligence

Ali Kashefi*
Department of Civil & Environmental Engineering
Stanford University, Stanford, CA 94305, USA
kashefi@stanford.edu

Abstract

In this technical report, we extensively investigate the accuracy of outputs from well-known generative artificial intelligence (AI) applications in response to prompts describing common fluid motion phenomena familiar to the fluid mechanics community. We examine a range of applications, including Midjourney, Dall·E, Runway ML, Microsoft Designer, Gemini, Meta AI, and Leonardo AI, introduced by prominent companies such as Google, OpenAI, Meta, and Microsoft. Our text prompts for generating images or videos include examples such as "Von Kármán vortex street", "flow past an airfoil", "Kelvin-Helmholtz instability", "shock waves on a sharp-nosed supersonic body", etc. We compare the images generated by these applications with real images from laboratory experiments and numerical software. Our findings indicate that these generative AI models are not adequately trained in fluid dynamics imagery, leading to potentially misleading outputs. Beyond text-to-image/video generation, we further explore the transition from image/video to text generation using these AI tools, aiming to investigate the accuracy of their descriptions of fluid motion phenomena. This report serves as a cautionary note for educators in academic institutions, highlighting the potential for these tools to mislead students. It also aims to inform researchers at these renowned companies, encouraging them to address this issue. We conjecture that a primary reason for this shortcoming is the limited access to copyright-protected fluid motion images from scientific journals.

Keywords AI-generated fluid imagery · AI in education · Prompt engineering · Midjourney · Dall·E · Meta AI · Gemini

1 Introduction and motivation

Generative AI models [Gozalo-Brizuela and Garrido-Merchan, 2023, Feuerriegel et al., 2024, Lim et al., 2023, Cao et al., 2023, Epstein et al., 2023, Kashefi and Mukerji, 2023, Brynjolfsson et al., 2023, Fui-Hoon Nah et al., 2023] are nowadays popular and are used in various applications with transforming text, images, video, and sounds based on diverse input prompts. A subset of generative models is Large Language Models (LLMs) [Xu et al., 2022, Hoffmann et al., 2022, Thirunavukarasu et al., 2023, Kasneci et al., 2023, Chang et al., 2023, Chen et al., 2021, Kojima et al., 2022, Yang et al., 2023, Touvron et al., 2023, Roziere et al., 2023, Çelen et al., 2024], which have seen substantial development from famous tech companies such as Google, Microsoft, Open AI, Midjourney, and Runway ML.

One specific task of these models is to generate images from text prompts. As researchers in the fluid mechanics and applied mathematics communities, we are interested in investigating the capability of these generative AI models to produce accurate images based on textual prompts describing fluid mechanics phenomena. For our investigation, we use common textual prompts that are almost universally known in the community for a fair assessment.

Image generation using textual prompts can be assessed on two levels. At the first level, the question is whether the generative AI models can generate at least a relevant image. A relevant image can be thought of as something that first comes to mind for a fluid mechanics expert. For instance, if the prompt is "Von-Karmann Vortex Street", [Tritton, 1959, Kashefi and Staples, 2018, Kashefi, 2020a, 2022] an image similar to the cover of the "An Album of Fluid Motion" book [Van Dyke and Van Dyke, 1982] might be expected. At a deeper level, the images should not only be relevant

*Corresponding author.

but also accurately describe the characteristic features of fluid motion. For example, for the prompt "Flow over a backward-facing step at Reynolds number of 400", [Kim and Moin, 1985, Armaly et al., 1983, Le et al., 1997, Kashefi and Staples, 2018, Kashefi, 2020b] it is crucial to assess whether the reattachment length in the generated image is correct. This technical report focuses on the first level. We compare generative images with those found in open-source articles from the Journal of Fluid Mechanics, which are derived from lab experiments or numerical software.

The remainder of this technical report is organized as follows. Sect. 2 offers a brief review of the governing equations of fluid motion along with two special cases of compressible and incompressible flow. Sect. 3, which is the main part of this report, focuses on text-to-image generation for fluid motion using well-known AI tools. In Sect. 4, an example of text-to-video generation relevant to fluid motion is provided. Image-to-text and video-to-text generation, as applied to fluid dynamic motion, are investigated in Sect. 5 and Sect. 6, respectively. Lastly, we provide a summary and concluding remarks of our technical report in Sect. 7.

2 Brief overview of fluid dynamic motion

In this subsection, we briefly review the governing equations for fluids at the most general level. The two main principles are the continuity equation and the Cauchy momentum equation [Euler, 1757, Łukaszewicz and Kalita, 2016, Constantin and Foiaş, 1988, Chorin et al., 1990, Batchelor, 1967] and can be written respectively as

$$\frac{D\rho}{Dt} + \rho\left(\nabla \cdot \boldsymbol{u}\right) = 0, \tag{1}$$

$$\rho\frac{D\boldsymbol{u}}{Dt} = \rho\boldsymbol{f} + \nabla \cdot \boldsymbol{T}, \tag{2}$$

where ρ represents the fluid density, and $\boldsymbol{u}$ represents the velocity vector. The total derivative with respect to time is denoted by $\frac{D}{Dt}$. The body force is represented by $\boldsymbol{f}$. $\boldsymbol{T}$ represents the most general form of the stress tensor. Under different assumptions, these two equations (Eqs. 1–2) can be simplified. For instance, in the case of compressible flow [Feistauer et al., 2003, Novotny and Straskraba, 2004], the stress tensor is expressed as

$$\boldsymbol{T} = -p\boldsymbol{I} + \lambda\left(\nabla \cdot \boldsymbol{u}\right)\boldsymbol{I} + \mu\left(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^T\right), \tag{3}$$

where p indicates the fluid pressure. Dynamic viscosity and the second viscosity are respectively shown by μ and λ. $\boldsymbol{I}$ is the identity tensor. As another simplification, in the case of incompressible flow [Majda et al., 2002, Panton, 2024], the velocity vector is divergence-free, which is mathematically represented as

$$\nabla \cdot \boldsymbol{u} = 0, \tag{4}$$

leading to further simplification of Eq. 3. These equations (Eqs. 1–2), which may be coupled with others, describe the behavior of fluid motion in an Eulerian system. Similarly, fluid motion can be described in a Lagrangian system. In a specified time-space domain, along with boundary and initial conditions, the equations can be solved numerically, and the results are illustrated using visualization software [Matsson, 2023, Squillacote et al., 2007, Ayachit, 2015, Yang et al., 2020, Jasak, 2009, Jasak et al., 2007, Heidbach et al., 2020, Pryor, 2009, Flaischlen and Wehinger, 2019, Mullen, 2011]. In this technical report, we consider images of fluid motions obtained by both experimental [Huang et al., 1997, Westerweel et al., 1997, Willert and Gharib, 1991, Van Dyke and Van Dyke, 1982, Adrian, 1991, Tropea et al., 2007, Raffel et al., 2018] and numerical tools [Hirsch, 2007, Patankar, 2018, Osher and Fedkiw, 2004, Anderson and Wendt, 1995, Reddy and Gartling, 2010].

3 Text to image

In this subsection, we first briefly introduce different generative applications and then compare their outputs for eleven different text prompts related to fluid dynamics motion. Specifically, we consider Midjourney [2] (Sect. 3.1), ChatGPT-4 [3] and Microsoft Designer [4] with DALL·E 3 (Sect. 3.2), Runway ML [5] (Sect. 3.3), Gemini [6] (Sect. 3.4), Meta AI [7] (Sect. 3.5), and Leonardo AI [8] (Sect. 3.6).

[2]https://www.midjourney.com/home
[3]https://chat.openai.com
[4]https://designer.microsoft.com
[5]https://runwayml.com/
[6]https://gemini.google.com/app
[7]https://ai.meta.com/meta-ai/
[8]https://leonardo.ai

3.1 Midjourney

Midjourney [Borji, 2022, Hanna, 2023, Huang et al., 2023, Henderson et al., 2023] was released in July 2022 by the research lab Midjourney. Its primary goal is to generate images from users' textual prompts. The platform offers different plans, with the basic one priced at $10 per month. Access to this AI platform is provided through Discord. Additionally, users can view the prompts and images generated by others. The capabilities of Midjourney have been investigated by other researchers in areas such as graphic design [Mansour, 2023], architecture [Jaruga-Rozdolska, 2022], industrial design [Yin et al., 2023], etc.

3.2 DALL·E (embedded in ChatGPT 4, ChatGPT 4o, and Microsoft Designer)

DALL·E was developed and released by OpenAI in 2021 [Kang and Yi, 2023, Leivada et al., 2023, Borji, 2022, Marcus et al., 2022, Betker et al., 2023, Adams et al., 2023, Hwang and Oh, 2023, Henderson et al., 2023, Ajmera et al., 2024]. The most recent version, DALL·E 3, is now integrated with ChatGPT 4, and recently ChatGPT 4o. As a result, users of ChatGPT 4 and ChatGPT 4o can generate images using textual prompts. Currently, the basic plan for ChatGPT 4 and ChatGPT 4o is priced at $20 per month, with a limit of 40 messages every three hours.

It is noted that Microsoft Designer currently uses DALL·E 3 for generating images as well. Microsoft Designer [Hwang and Oh, 2023, Henderson et al., 2023, Pearson, 2023] was launched by Microsoft and is currently still in its experimental phase. This AI tool generates photos based on user-generated textual prompts, and the platform is currently available for free to users.

3.3 Runway ML

Runway ML [Hales, 2021, Othman, 2023, Gupta, 2021] was launched in 2018 by Runway AI, Inc. This platform has a variety of applications, but we specifically explore its capabilities in text-to-image and text-to-video generation, particularly in the context of fluid motion. The current version of Runway operates based on the deep learning model of Stable Diffusion [Rombach et al., 2022, Esser et al., 2024, Meng et al., 2023, Ho et al., 2022]. Notably, researchers from Runway AI, Inc. contributed to the development of the Stable Diffusion algorithm [Borji, 2022]. It is worth noting that Stability AI is another notable company that contributed to the development of the Stable Diffusion algorithm [Esser et al., 2024]. However, in this work, our examination is focused on Runway ML. Note that Runway ML offers different styles for generating images. In the current article, we use the *None* style.

3.4 Gemini Advanced

Gemini, a large language model developed by Google DeepMind, originally known as Bard, can generate images from textual descriptions and describe image content in text [Masalkhi et al., 2024, McIntosh et al., 2023, Singh et al., 2023, Aydin, 2023, Campesato, 2024, Carlà et al., 2024, Buscemi and Proverbio, 2024, Lu et al., 2024]. This report explores both capabilities of Gemini: text-to-image in this subsection and image-to-text in Sect. 5.3. The basic version of Gemini is free, while its advanced version is free for two months and then costs $20 per month. We use the advanced version (i.e., Gemini Advanced) for the current study.

3.5 Meta AI (built on Meta LLaMA 3)

Meta AI, released by Meta, is a large language model built on the Meta LLaMA 3 architecture [Touvron et al., 2023, Bhatt et al., 2024, Li et al., 2024a, Harper, 2024]. The model is currently available for free, but only in specific countries. The model has the capability to generate images based on text prompts.

3.6 Leonardo Ai

Leonardo Ai is another tool that operates using deep learning models [Ruiz-Rojas et al., 2023, Wang and Utama, 2023, Moreno-Sánchez et al., 2023]. For the purposes of this technical report, we examine its capabilities for text-to-image generation. Additionally, we utilize the paid version of the service ($12 per month) for this exploration. Note that Leonardo Ai offers different models with different modes. In this study, we use the model of "Leonardo Lightning XL" with the "Dynamics" mode.

(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai

(g) Von Kármán vortex street behind an oscillating circular cylinder in a lab experimental setup; this image is taken from part (c) of Fig. 25 in Khan et al. [2023], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 1: A comparison between the AI-generated images and the lap experimental result (i.e., ground truth) for the prompt "Von Kármán vortex street"

3.7 Comparison

In this subsection, we compare the performance of these six different generative models. As a general structure for this subsection, we present images generated by these generative models, along with a sub-figure in part (g) of Figs. 1–11 that shows an accurate image (as ground truth) associated with the prompt relevant to that fluid dynamics phenomenon.

3.7.1 Von Kármán vortex street

The first textual prompt we investigate in this subsection is "Von Kármán vortex street". Before analyzing the outputs from the generative models, we highlight a few aspects of this fluid dynamics problem, specifically by mentioning the Reynolds number (Re). An important dimensionless number in fluid mechanics is the Reynolds number (see e.g., [Rott, 1990, Johnson and Patel, 1999, Purcell, 2014, Smits et al., 2011, Drela, 1989, Kashefi et al., 2021, Kashefi, 2020c]). Depending on the context of the problem, it can be written in different forms. For instance, for the problem of flow past a circular cylinder, which can be seen in part (g) of Fig. 1 and part (g) of Fig. 2, it can be expressed as follows:

$$Re = \frac{\rho d U_\infty}{\mu}, \tag{5}$$

where ρ and μ are respectively the density and the dynamic viscosity of the fluid. The free stream velocity and the cylinder diameter are shown by U_∞ and d, respectively. The Reynolds number (Re) is used to characterize other important fluid motion problems, such as flow over an airfoil (see part (g) of Fig. 8), flow over an airplane wing (see part (g) of Fig. 11), and Kelvin–Helmholtz instability (see part (g) of Fig. 4), among others.

One of the dimensionless numbers relevant to the Von Kármán vortex street (see e.g., [Von Kármán, 2004]), which can be seen in part (g) of Fig. 1, is the Strouhal number (see e.g., [Okajima, 1982]), which is formulated as follows:

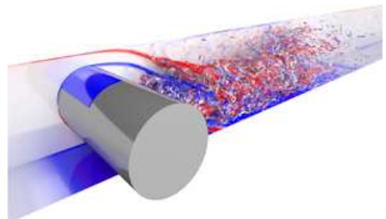
$$St = \frac{f_s d}{u_\infty}, \tag{6}$$

where f_s represents the shedding frequency, d is the cylinder diameter similar to what we define in Eq. 5, and u_∞ is the free stream velocity of the fluid.

Figure 1 depicts responses to the prompt "Von Kármán vortex street". In images generated by Midjourney and Leonardo Ai, a street is visible, indicating that these generative models are misled by the term "street" in the prompt. The concept of "vortex" is observed in images generated by Midjourney, Gemini Advanced, DALL·E 3, Meta AI, and Runway; however, none of these representations acceptably depict the vortex in a Von Kármán vortex street.

The image generated by Meta AI is interesting in that it depicts water circulation in a pipe within a laboratory environment. Although this image does not represent the Von Kármán vortex street, it is relevant to fluid mechanics.

(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai



(g) Flow past a three-dimensional circular cylinder at a Reynolds number of $Re = 3900$; this image, showing vorticity and produced by a numerical simulation, serves as the abstract photo in Lu et al. [2023], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 2: A comparison between the AI-generated images and the numerical simulation result (i.e., ground truth) for the prompt "Flow past a circular cylinder"

Interestingly, the image generated by DALL·E 3 could be considered relevant to the rotating disk problem or the Von Kármán swirling flow problem [Cochran, 1934, Kármán, 1921].

As a general observation, it seems that these generative models do not interpret "Von Kármán vortex street" as a cohesive expression. Instead, they treat it as a combination of the individual words "vortex" and "street", attempting to generate images that merge these two concepts.

3.7.2 Flow past a circular cylinder

Figure 2 presents images generated in response to the prompt "flow past a circular cylinder". None of the generative models under investigation accurately captures the phenomenon. Among the generated images, the result from Midjourney is notable as it features fluid rotating in a hollow cylinder.

In the fluid mechanics community, the term "flow" typically implies fluid flow, such as water or air. Bearing this in mind, although some types of flow are visible in the images generated by DALL·E 3 and Meta AI, they do not depict actual fluid flow, perhaps due to the fact that the term "fluid" is not included in the prompt "flow past a circular cylinder". In all of these generated images, we can observe the presence of a cylinder as an object. Interestingly, each generative model depicts the cylinder from a unique perspective and angle.

The image generated by Gemini Advanced is interesting because, although it is not directly relevant to "flow past a circular cylinder", it depicts a schematic of flow inside an inclined cylinder. Additionally, this image includes some written letters. This type of figure can often be found in fluid mechanics textbooks (see, for example, Fig. P5.131 of [Munson et al., 2013]). A similar observation applies to the image generated by Runway, which can be interpreted as showing the cross-section of a cylinder. The figure generated by Leonardo Ai is remarkable because it exhibits flow evolving on the outer surface of a cylinder, perhaps the closest representation to what we expected, yet it remains significantly distant from the actual phenomenon.

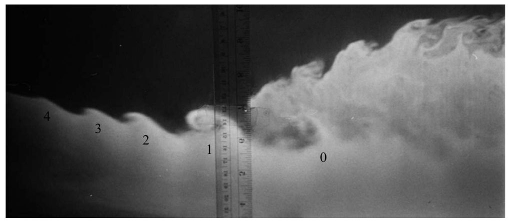
3.7.3 Hydraulic jump

The next prompt considered in this technical report is "hydraulic jump", with the corresponding images displayed in Fig. 3. The occurrence of the hydraulic jump, as observed in part (g) of Fig. 3 and part (g) of Fig. 9, can be identified using the dimensionless Froude number (see e.g., [Hornung et al., 1995]), defined as:

$$Fr = \frac{U_u^2}{g h_u}, \tag{7}$$

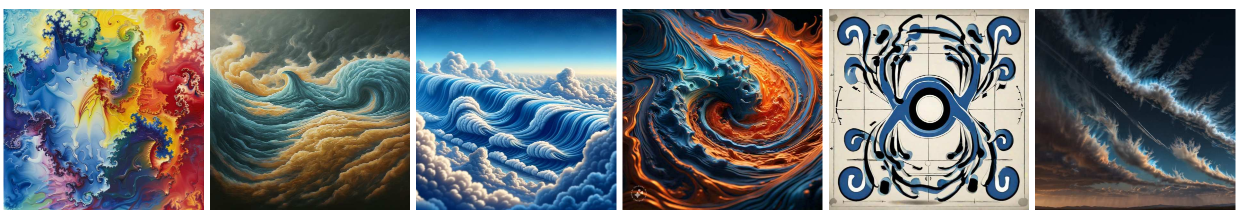
where U_u and h_u respectively represent the upstream velocity and height of the fluid. For $Fr > 1$, a hydraulic jump is expected.

(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai

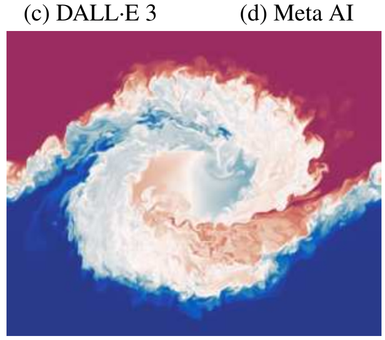


(g) Hydraulic jump in a lab experimental setup; a series of Kelvin–Helmholtz billows is also observable. This image is taken from part (b) of Fig. 9 in Lawrence and Armi [2022], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 3: A comparison between the AI-generated images and the lap experimental result (i.e., ground truth) for the prompt "Hydraulic jump"



(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai



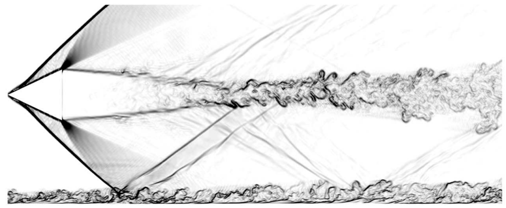
(g) Kelvin–Helmholtz instability; this image, showing a density contour and produced by a numerical simulation, serves as the abstract photo in Lewin and Caulfield [2022], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to freely distribute the content.

Figure 4: A comparison between the AI-generated images and the numerical simulation result (i.e., ground truth) for the prompt "Kelvin–Helmholtz instability"

As can be seen from Fig. 3, the images generated by Midjourney, Runway, and Leonardo Ai are inaccurate. Similar to some results discussed for previous prompts, these generative models interpret "hydraulic jump" as two separate words, "hydraulic" and "jump". Consequently, they depict unrelated objects such as motorcycles, minibus, or even humans jumping over a river.

On the other hand, the results by Gemini Advanced and Meta AI could be considered moderately relevant. These two generative models produce images that are partially accurate, as they depict water exiting a dam and a step in a river, two scenarios in nature where a hydraulic jump can occur, based on our knowledge. The image generated by DALL·E 3, however, does not seem informative and is difficult to judge.

(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai



(g) Shock wave around a cone at a Mach number of $M_\infty = 2.05$; This image, produced by a numerical simulation, is taken from the top part of Fig. 4 in Zuo et al. [2023], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 5: A comparison between the AI-generated images and the numerical simulation result (i.e., ground truth) for the prompt "shock waves on a sharp-nosed supersonic body"

3.7.4 Kelvin-Helmholtz instability

Figure 4 displays the collective images generated in response to the prompt "Kelvin-Helmholtz instability". Surprisingly, all of the images are relevant to this phenomenon. Images generated by Gemini Advanced, DALL·E 3, and Leonardo Ai attempt to depict the Kelvin-Helmholtz instability in clouds in the sky, which is a common manifestation of this instability in nature (see e.g., [Baumgarten and Fritts, 2014]). Images produced by Midjourney, Meta AI, and Runway depict the mixing of different liquids with varying colors, a phenomenon characteristic of Kelvin-Helmholtz instability, as shown in part (g) of Fig. 4, which serves as a ground truth example. It is worth mentioning that the images generated by Runway differ slightly from the others in terms of context.

3.7.5 Shock waves on a sharp-nosed supersonic body

Figure 5 exhibits a collection of images generated in response to the prompt "shock waves on a sharp-nosed supersonic body". A shock can be categorized using the dimensionless number known as the Mach number (see e.g., [Davis and Chyu, 1966, Howe, 1979]), formulated as:

$$M_\infty = \frac{V_\infty}{c}, \tag{8}$$

where V_∞ represents the free stream velocity and c is the sound speed. A Mach number of $M_\infty = 2.05$ indicates a supersonic flow accompanied by shock waves, as illustrated in part (g) of Fig. 5.

Based on observations from Fig. 5, the images produced by Gemini Advanced and DALL·E 3 can be considered the most relevant, especially when compared to a ground truth figure shown in part (g) of Fig. 5. The picture produced by Leonardo Ai is also interesting as it depicts shock boundaries, although it lacks a sharp-nosed body. The result by Runway is difficult to judge. It illustrates an educational poster with text printed below that reads, "Shock Waes aver Srpeouoc Edoy", closely mirroring the prompt "shock waves on a sharp-nosed supersonic body". Midjourney and Meta AI generate images that feature warplanes as sharp-nosed supersonic bodies, representing shock waves in an artistic rather than a scientific manner.

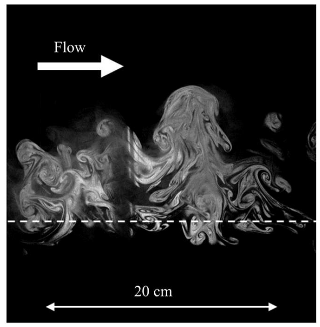
3.7.6 Rayleigh-Taylor instability

The next prompt is "Rayleigh-Taylor instability", with the generated images displayed in Fig. 6. An important dimensionless number for investigating Rayleigh-Taylor instability, which can be observed in Fig. 6, is the Atwood number (see e.g., [Luo et al., 2020, Ker, 1988, Andrews and Dalziel, 2010]), expressed as:

$$A = \frac{\rho_1 - \rho_2}{\rho_1 + \rho_2}, \tag{9}$$

where ρ_1 and ρ_2 are the densities of the heavier and lighter fluids, respectively.

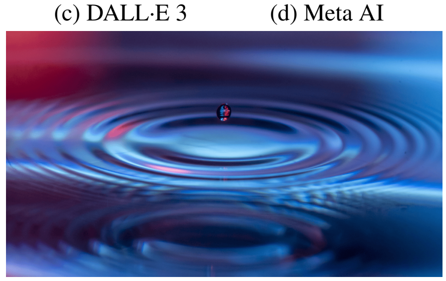
(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai



(g) Rayleigh-Taylor instability in a lap experimental setup; the picture is taken from part (a) of Fig. 5 in Suchandra and Ranjan [2023], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 6: A comparison between the AI-generated images and the lap experimental result (i.e., ground truth) for the prompt "Rayleigh-Taylor instability"

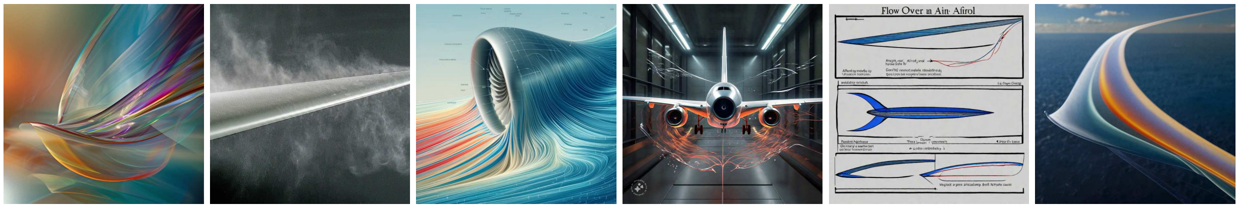
(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai



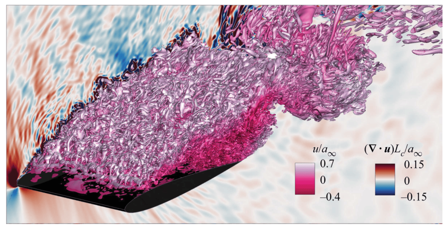
(g) Capillary wave: A small water droplet, approximately 0.4 mm in radius, rebounds from the same fluid bath in a lab experimental setup; this image is Fig. 1 in Alventosa et al. [2023], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 7: A comparison between the AI-generated images and the numerical simulation result (i.e., ground truth) for the prompt "capillary wave"

According to observations from Fig. 6, the image produced by DALL·E 3 is particularly interesting as it suggests the interaction between heavy and light fluids. The result from Gemini Advanced is noteworthy because it displays the interaction of two fluids. The outputs from Midjourney and Leonardo Ai are somewhat similar to the images generated for the "Kelvin-Helmholtz instability" prompt, as shown in Fig. 4. Conversely, the outputs from Meta AI and Runway are not accurate and largely irrelevant.



(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai



(g) Streamwise vorticity over an airfoil; this image, produced by a numerical simulation, is taken from part (a) of Fig. 1 in Turner and Kim [2022], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 8: A comparison between the AI-generated images and the numerical simulation result (i.e., ground truth) for the prompt "Flow over an airfoil"

3.7.7 Capillary wave

"Capillary wave" is the next prompt, and the generated images are collected in Fig. 7. The most accurate result is generated by Gemini Advanced. One possible question is whether Gemini Advanced uses online information simultaneously for generating images or if it only uses a predefined dataset. Following Gemini Advanced, the output from Meta AI is also relatively accurate and contextually relevant. In contrast, the images generated by Midjourney, Runway, and Leonardo Ai are misleading, as they depict general waves, such as sea waves. Meanwhile, the result from DALL·E 3, while artistic, does not pertain to capillary waves.

3.7.8 Flow over an airfoil

In Fig. 8, we present a collection of generated images for the prompt "Flow over an airfoil". Gemini Advanced produces the most relevant image and it depicts an airplane wing. This wing might resemble NACA airfoils [Abbott et al., 1945], where the cross-section of the wing can be considered as an airfoil.

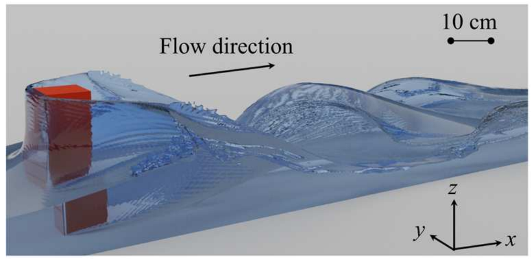
The image generated by DALL·E 3 depicts an airplane turbine, while Meta AI produces an image of an airplane in a room. The image generated by Runway is conceptually similar to those produced by Runway in Fig. 2 and Fig. 5. Additionally, the letters "Flow over m Ain Afirol" can be seen in that image. Images generated by Midjourney and Leonardo Ai are artistically similar to each other, yet they do not accurately represent the fluid mechanics concept of flow over an airfoil. Similar to our discussion on the prompt "Flow past a circular cylinder" in Sect. 3.7.2, the term "flow" may not be understood as "fluid flow" by these generative models.

3.7.9 Free-surface channel flow

The images generated in response to the prompt "Free-surface channel flow" are displayed in Fig. 9. This prompt is intriguing because it is a common term in everyday life and simultaneously a scientific term in fluid dynamics. Interestingly, this phenomenon shares features in both contexts. As illustrated in Fig. 9, all the generated images depict free surface flow in nature, resembling everyday photos. Conversely, part (g) of Fig. 9 presents a scientific picture associated with this phenomenon.

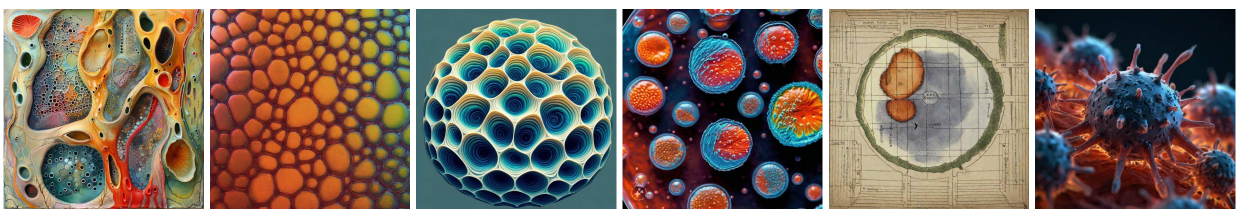
This simple experiment with prompts suggests the idea of allowing users to choose the context in generative models, particularly those available as applications. This feature would be beneficial when a prompt is common across different contexts or domains of knowledge, enabling users to select the specific context for which they want to generate images. For instance, Leonardo Ai currently allows users to specify a domain of interest for generating images, such as cinematic, environment, general, photography, etc. However, an academic option is not yet available.

(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai

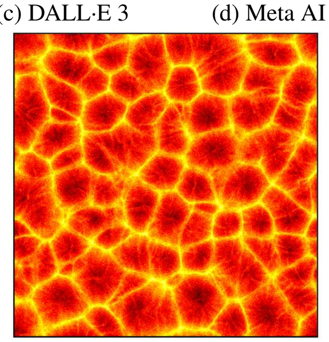


(g) Free-surface flow in a channel around a square cylinder at a Froude number of $Fr = 2.5$; a hydraulic jump is observed. The picture, produced by a numerical simulation, is taken from part (c) of Fig. 3 in Eames and Robinson [2024], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 9: A comparison between the AI-generated images and the numerical simulation result (i.e., ground truth) for the prompt "Free-surface flow"



(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai



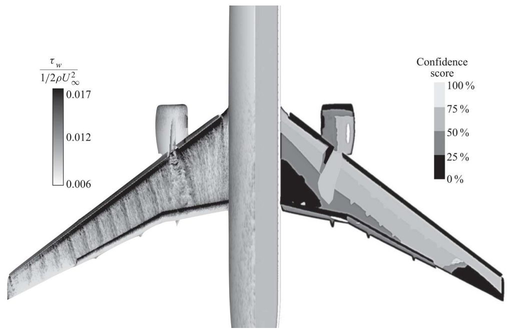
(g) Bénard cells resulting from Rayleigh–Bénard convection in a porous medium. This image, produced by a numerical simulation, is taken from part (b) of Fig. 9 in De Paoli et al. [2022], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 10: A comparison between the AI-generated images and the numerical simulation result (i.e., ground truth) for the prompt "Bénard cells"

3.7.10 Bénard cells

In Fig. 10, we examine the images generated by the models in response to the prompt "Bénard cells". Except for the image produced by Gemini Advanced, the rest are irrelevant. Similar to what we discussed in the previous part, this accuracy from Gemini Advanced may suggest it has access to online information. The other models display images of cells, likely reflecting contexts such as biology or medicine, indicating a misunderstanding of "Bénard cells" as a technical term in fluid mechanics.

(a) Midjourney (b) Gemini (c) DALL·E 3 (d) Meta AI (e) Runway (f) Leonardo Ai



(g) Normalized wall shear stress (left) and confidence map (right) displayed over the NASA Common Research Model High-lift (`https://commonresearchmodel.larc.nasa.gov/high-lift-crm/`). This image, produced by a numerical simulation, is Fig. 16 in Lozano-Durán and Bae [2023], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Figure 11: A comparison between the AI-generated images and the numerical simulation result (i.e., ground truth) for the prompt "Flow over an airplane wing"

3.7.11 Flow over an airplane wing

The last prompt investigated in this technical report is "Flow over an airplane wing", with the outputs from the generative models displayed in Fig. 11. Midjourney, Gemini Advanced, Runway, and Leonardo Ai each generate an image of an airplane wing. These images are very similar to each other, displaying a similar perspective and angle of view. They depict wings of airplanes flying in the sky above clouds. However, it is not clear how the term "flow" in our textual prompt is reflected in these images, as can be seen in Fig. 11.

According to Fig. 11, the generated image by Meta AI is notable because it shows an airplane moving downward, with water flowing around the wing. The photo appears artistic; however, it can be considered an alternative interpretation of the textual prompt. Indeed, although we did not initially expect to see this image in response to the prompt "flow over an airplane wing", the image effectively illustrates the action of an object falling into water, capturing a realistic sense in the context of fluid dynamics.

From what we observe in Fig. 11, the image generated by DALL·E 3 is moderately similar to what DALL·E 3 produced for the prompt "flow over an airfoil", shown in Fig. 8. Again, we see a turbine attached to an airplane wing. However, this time the wing is depicted in a very cinematic style.

4 Text to video

4.1 Meta AI

In Sect. 3.5, we used Meta AI to generate images of fluid motion from text descriptions. In this subsection, we investigate Meta AI's capability to generate videos from text prompts. It should be noted that Meta AI does not directly create videos from text prompts. Instead, it first generates images, and then users can animate these images by pressing the "animate" button. Meta AI automatically animates the images, producing animations approximately three seconds in length.

4.2 Runway ML

In Sect. 5.2, we used Runway ML for generating images from text, specifically depicting fluid dynamic motion. In this subsection, we shift to exploring Runway ML's text-to-video capabilities, focusing on creating short videos that visually represent fluid dynamic motion based on textual descriptions. Similar to Meta AI, Runway ML first generates a picture and then animates it. However, compared to Meta AI, Runway ML produces images of higher quality.

4.3 Comparison

To compare the performance of Meta AI and Runway ML, we use the textual prompt “water past a circular cylinder”. The generated videos are available in the repository mentioned in the data availability section. To illustrate the outputs, we display several sequential frames from the generated videos by Meta AI and Runway ML respectively in Fig. 12 and Fig. 13. The result from Runway ML could be aesthetically valuable. It appears that the cylinder is fixed while water flows around it before entering a river. The interaction between this water and the river water is interesting, as it creates waves. These waves could be considered similar to capillary waves. The image generated by Meta AI is striking, as it appears that the cylinder rotates, causing the water inside to also rotate. Furthermore, the cylinder is partially submerged in pond water, and the interaction between the rotating cylinder and the water surface is displayed. Overall, the generated videos do not deliver the expected and relevant results for the prompt “water past a circular cylinder”; however, they do generate other interesting fluid dynamics phenomena. Finally, it should be noted that the reason for using the term “water” instead of “flow” or “fluid flow” is intentional, as these generative models have issues interpreting “flow” in the context of fluid dynamics. This issue was discussed in detail in Sect. 3.7.

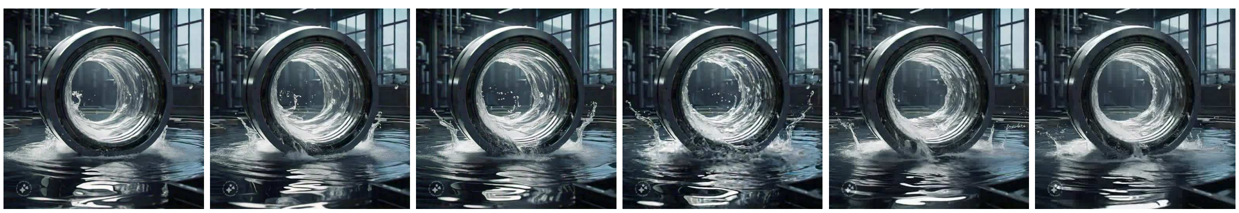


Figure 12: Several sequential frames from the generated video by Meta AI for the prompt “water past a circular cylinder”. The full generated video is available at this GitHub repository: `https://github.com/Ali-Stanford/MisleadingGalleryOfFluidMotionByAI`.

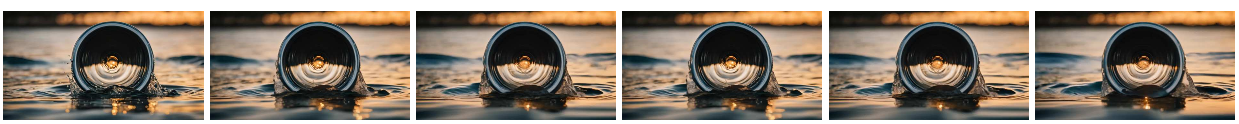


Figure 13: Several sequential frames from the generated video by Runway ML for the prompt “water past a circular cylinder”. The full generated video is available at this GitHub repository: `https://github.com/Ali-Stanford/MisleadingGalleryOfFluidMotionByAI`.

5 Image to text

5.1 LLaVA

LLaVA [9] was primarily developed by researchers at Microsoft and the University of Wisconsin-Madison [Liu et al., 2023a, Lin et al., 2023, Buscemi and Proverbio, 2024]. It is designed to generate textual descriptions from input images and videos. LLaVA is open-source and free to use, and it has been widely applied in various fields such as biomedicine [Li et al., 2023], AI agents [Liu et al., 2023b], and solving mathematical geometric problems [Gao et al., 2023]. In this subsection, we utilize image-to-text generation, while in Sect. 6.1, we employ the video-to-text generation capability of LLaVA.

5.2 ChatGPT 4

Beyond the capability of generating images from textual descriptions, ChatGPT 4 is also capable of describing the content of a given image in text [Ren et al., 2024, Lu et al., 2024, Zhu et al., 2023, Johnson et al., 2023]. In this subsection, we analyze this aspect of ChatGPT 4, focusing on fluid dynamic motion images.

[9] https://llava.hliu.cc

5.3 Gemini

In Sect. 3.5, we explored Gemini's capability for generating textual descriptions of images depicting fluid motion. In this subsection, we assess Gemini's performance in providing textual descriptions for an image illustrating fluid motion.

5.4 Comparison

In this section, we explore the capabilities of three different generative models: ChatGPT-4, Gemini Advanced, and LLaVA, in generating textual descriptions of images depicting fluid motion. We upload a photo of fluid motion to these applications and then ask each to describe the image. To compare these models, we consider two different images. The first image is the Von Karman vortex street shown in part (g) of Fig. 1. The generated textual descriptions are provided in the following box.

ChatGPT-4

The image appears to capture a fluid dynamic phenomenon known as a Karman vortex street. This occurs when a fluid flows around a cylindrical object, causing the flow to separate and form vortices on the downstream side. The photograph shows several swirling patterns with bright yellow lines on a dark background, indicating the path of the fluid flow and the vortices created. The yellow lines could be a dye or some form of visual aid to show the flow pattern in the fluid, often used in fluid dynamics studies to visualize the motion. There's a sequence of alternating vortices that create a somewhat symmetric pattern across the image. The overall aesthetic is quite artistic and mesmerizing, capturing both the beauty and complexity of fluid motion.

Gemini Advanced

The image you sent me is a close-up of a lemon peel. The peel is a bright yellow color, and it has a rough, textured surface. The black background makes the yellow of the peel stand out. There is a small white letter "C" in the bottom right corner of the image.

LLaVA

The image appears to be a close-up photograph of a substance with a flowing, swirling pattern. The substance has a yellowish hue and is captured in motion, creating a dynamic, abstract shape that resembles a snake or a ribbon. The background is dark, which contrasts with the bright yellow of the substance, making the swirling pattern stand out. The image has a shallow depth of field, with the focus on the swirling substance, causing the background to be blurred. There is a letter "C" in the top right corner, which might be a watermark or a part of a logo. The overall effect is quite mesmerizing and could be interpreted as a representation of movement or fluid dynamics.

As indicated in the above box, ChatGPT-4 provides the best description compared to Gemini Advanced and LLaVA. ChatGPT-4 clearly mentions the "Karman vortex street" and describes the phenomenon by addressing keywords such as "vortex" and "flow past a cylinder". It also recognizes the colors yellow and black in the image. More importantly, ChatGPT-4 notes the yellow color as indicative of an experimental technique used to visualize the vortex.

While the LLaVA description is not entirely accurate, it contains significantly more accurate information compared to Gemini Advanced. For instance, LLaVA addresses the swirling pattern and includes terms such as "movement of fluid dynamics" in its description. On the other hand, the Gemini Advanced response is almost unrelated. Interestingly, both Gemini Advanced and LLaVA refer to the letter "C" in the picture. Note that as mentioned earlier, LLaVA is free, whereas Gemini Advanced costs $12 per month.

Another interesting and important point is that when we ask DALL·E 3, embedded in ChatGPT-4, to generate a picture of the Von Kármán vortex street in Sect. 3, the result is inaccurate, as exhibited in part (c) of Fig. 1. However, it provides an accurate textual description when it analyzes an actual picture of the Von Kármán vortex street.

The second image depicts another relatively famous phenomenon in fluid dynamics: the corona splash, which occurs at the moment a droplet impacts a liquid surface, as shown in Fig. 14. The textual descriptions by ChatGPT-4, Gemini Advanced, and LLaVA are provided in the following box.

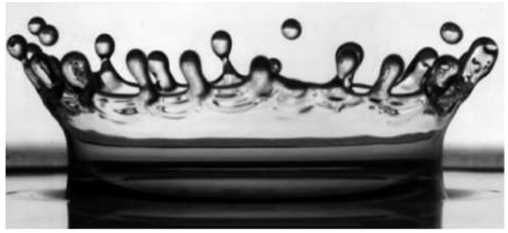


Figure 14: Corona splash in a lab experimental setup; this image is taken from part (c) of Fig. 2 in Burzynski et al. [2020], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

ChatGPT-4

This image captures a moment of fluid dynamics beautifully. It's a black and white photograph showing a crown-like splash of water. The base of the splash is broad and relatively calm, which suggests that it is probably the surface of water in a container. From this base, several columns of water rise and arc outwards, with droplets separating from the tips of these arcs, forming a pattern that resembles a tiara or a crown. The splash is symmetric, and the physics behind this phenomenon involve the surface tension of the water and the forces exerted upon the initial impact, likely from a droplet falling into the water. The clarity and sharpness of the image suggest a high-speed photograph to freeze the motion.

Gemini Advanced

The image you sent is a black and white photo of a crown made of water droplets. The water droplets are arranged in a circular pattern, with some droplets larger than others. The background is dark and blurry. The overall effect is one of fragility and impermanence, as the water droplets appear to be on the verge of falling apart.

LLaVA

The image shows a glass of water with water droplets being poured into it. The droplets are captured in mid-air, creating a dynamic and splashy effect. The water appears to be clear, and the glass is transparent, allowing us to see the water inside. The background is a plain, light color, which contrasts with the dark tones of the water and the droplets, emphasizing the action of the pouring. The image is in black and white, which adds a timeless and artistic quality to the scene.

Similar to the previous investigation, and now focusing on the image shown in Fig. 14, the output from ChatGPT-4 is significantly more accurate and informative than those generated by Gemini Advanced and LLaVA. ChatGPT-4 clearly mentions the term "splash". More interestingly, it addresses the physics of surface tension and accurately recognizes that the photo was taken with a high-speed camera. Both Gemini Advanced and LLaVA describe the water splashing and droplets, and their results are more relevant compared to the image of the Von Karman vortex street.

6 Video to text

6.1 Video-LLaMA

In this subsection, we evaluate the performance of Video-LLaMA [10] [Zhang et al., 2023] in providing textual descriptions for a short movie that shows fluid motion. We consider two short movies, both of which are taken from laboratory experimental studies published in the Journal of Fluid Mechanics. These movies are available as supplementary materials on the website of each journal paper. Therefore, potential audiences can refer to these websites to watch the movies. Additionally, we have included frames from the movies in Figures 13 and 14, as presented in the associated journal papers. The first movie depicts breaking waves, as illustrated by the frames shown in Fig. 15. The accompanying textual description from Video-LLaMA is provided in the box below.

[10] https://huggingface.co/spaces/DAMO-NLP-SG/Video-LLaMA

Figure 15: Breaking wave in a lab experimental setup: The corresponding movie is available as a "supplementary movie" at `https://doi.org/10.1017/jfm.2021.841`. This image is Fig. 3 in Steer et al. [2021], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

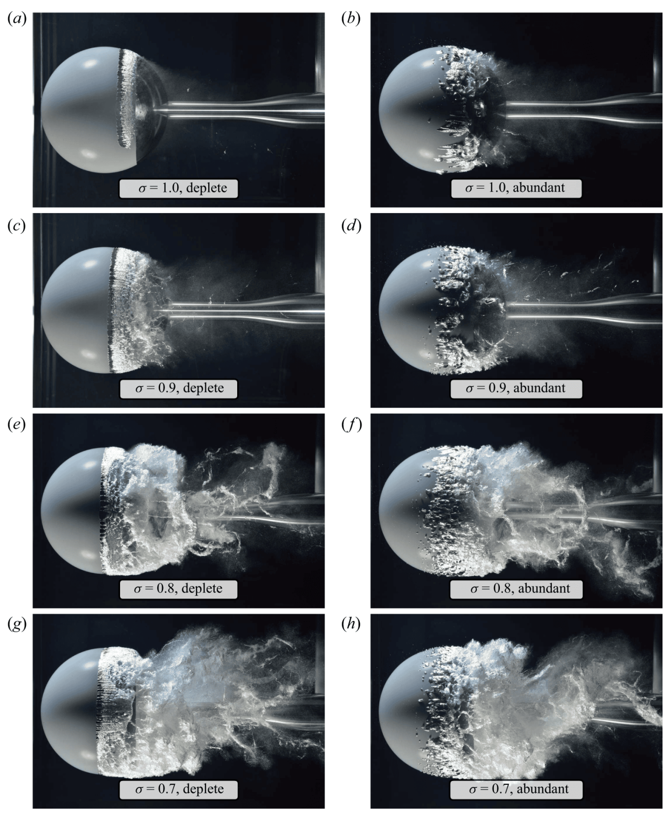


Figure 16: Drops and bubbles around a sphere showing nucleation effects on cavitation in a lab experimental setup: The corresponding movie is available as "supplementary movie 2" at `https://doi.org/10.1017/jfm.2022.511`. This image is taken from parts (a)–(h) of Fig. 6 in Brandner et al. [2022], published as open access under a Creative Commons license in the Journal of Fluid Mechanics, allowing readers to distribute the content freely.

Video-LLaMA
In this video, we see a black and white photograph of a room with water pouring on a table or floor. We see a variety of objects in the room, including a table, glasses, and a toilet. There is also a skateboard in the room and a clock in the background. We see different angles and views of the pouring water as it covers the table or floor.

The only accurate information in the textual description by Video-LLaMA is relevant to water being poured on a table or the floor. Video-LLaMA also attempts to identify objects, and the only accurately identified object is a table. However, it fails to address the main phenomenon depicted in the movie, which is wave breaking.

The second movie depicts drops and bubbles around a sphere, and some of its frames can be observed in Fig. 16. The text describing the movie, generated by Video-LLaMA, can be read as follows.

Video-LLaMA
The video consists of a black background with various objects and actions happening around it. The objects in the video include a metal ball on a table, a metal plate, a black liquid, a bottle of water, a cup, and a black mirror. The actions in the video include the presence of smoke, water on a mirror, and various objects moving.

Similar to the previous examination, Video-LLaMA cannot correctly predict and describe the phenomenon occurring in the movie. The only correct information it contains pertains to the presence of a ball and water. It seems that Video-LLaMA views the movies as a sequence of pictures; however, it is unable to make acceptable reasoning among the frames.

7 Summary and conclusion

This technical report comprehensively evaluated the capabilities of several generative AI models to generate text-to-image, text-to-video, image-to-text, and video-to-text in the context of fluid dynamics. Specifically, we investigated prominent models such as Midjourney, Dall·E3, Gemini Advanced, Meta AI, Runway ML, Leonardo Ai, Video-LLaMA, and LLaVA. Our findings demonstrated that the outputs generated by these AI models generally exhibited discrepancies when compared to realistic images obtained from numerical simulations or laboratory experiments, where fluid dynamics events are scientifically and accurately represented. The limitations of these generative models could be concerning in educational contexts, where they might potentially be used with students. Therefore, it is crucial to be aware of these models' shortcomings in handling domain-specific knowledge.

Our conjecture is that these limitations arise because these generative models have not been trained or refined with empirical data on fluid dynamics phenomena. This issue likely stems from the fact that such data, including images and videos, are often copyrighted by scientific journals and publishers, making them not publicly available. As a potential solution, educational institutions could request that the corresponding companies customize and refine these generative models specifically for fluid dynamics by providing sufficient data. We hope that this technical report serves as a starting point for further collaboration between AI companies developing these models and domain-specific experts in the fluid mechanics community to bridge the existing gap.

Finally, it should be noted that similar assessments have been conducted in other scientific disciplines for generative models (see e.g., [Meskó and Topol, 2023, Megahed et al., 2024, Theis et al., 2015, Yang and Lerch, 2020]). As a future direction, such investigations could be extended to additional areas of mechanical engineering and engineering mechanics [Xue et al., 2020, Regenwetter et al., 2022, Kashefi and Mukerji, 2022, Kashefi et al., 2023, Pal et al., 2024, Ni and Buehler, 2024, Polverini and Gregorcic, 2024, Luu and Buehler, 2024, Li et al., 2024b]. This expansion would provide further insights into the capabilities and limitations of these generative AI models across various engineering disciplines.

Data availability

The generated videos and high-quality images are available at the following GitHub repository: `https://github.com/Ali-Stanford/MisleadingGalleryOfFluidMotionByAI`.

References

Roberto Gozalo-Brizuela and Eduardo C Garrido-Merchan. Chatgpt is not all you need. a state of the art review of large generative ai models. *arXiv preprint arXiv:2301.04655*, 2023.

Stefan Feuerriegel, Jochen Hartmann, Christian Janiesch, and Patrick Zschech. Generative ai. *Business & Information Systems Engineering*, 66(1):111–126, 2024.

Weng Marc Lim, Asanka Gunasekara, Jessica Leigh Pallant, Jason Ian Pallant, and Ekaterina Pechenkina. Generative ai and the future of education: Ragnarök or reformation? a paradoxical perspective from management educators. *The international journal of management education*, 21(2):100790, 2023.

Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S Yu, and Lichao Sun. A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. *arXiv preprint arXiv:2303.04226*, 2023.

Ziv Epstein, Aaron Hertzmann, Investigators of Human Creativity, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R Frank, Matthew Groh, Laura Herman, Neil Leach, et al. Art and the science of generative ai. *Science*, 380(6650): 1110–1111, 2023.

Ali Kashefi and Tapan Mukerji. Chatgpt for programming numerical methods. *Journal of Machine Learning for Modeling and Computing*, 4(2), 2023.

Erik Brynjolfsson, Danielle Li, and Lindsey R Raymond. Generative ai at work. Technical report, National Bureau of Economic Research, 2023.

Fiona Fui-Hoon Nah, Ruilin Zheng, Jingyuan Cai, Keng Siau, and Langtao Chen. Generative ai and chatgpt: Applications, challenges, and ai-human collaboration, 2023.

Frank F Xu, Uri Alon, Graham Neubig, and Vincent Josua Hellendoorn. A systematic evaluation of large language models of code. In *Proceedings of the 6th ACM SIGPLAN International Symposium on Machine Programming*, pages 1–10, 2022.

Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch, Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, et al. Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556*, 2022.

Arun James Thirunavukarasu, Darren Shu Jeng Ting, Kabilan Elangovan, Laura Gutierrez, Ting Fang Tan, and Daniel Shu Wei Ting. Large language models in medicine. *Nature medicine*, 29(8):1930–1940, 2023.

Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. Chatgpt for good? on opportunities and challenges of large language models for education. *Learning and individual differences*, 103:102274, 2023.

Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, et al. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 2023.

Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, et al. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*, 2021.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. Large language models are zero-shot reasoners. *Advances in neural information processing systems*, 35:22199–22213, 2022.

Liu Yang, Siting Liu, Tingwei Meng, and Stanley J Osher. In-context operator learning with data prompts for differential equation problems. *Proceedings of the National Academy of Sciences*, 120(39):e2310142120, 2023.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

Baptiste Roziere, Jonas Gehring, Fabian Gloeckle, Sten Sootla, Itai Gat, Xiaoqing Ellen Tan, Yossi Adi, Jingyu Liu, Tal Remez, Jérémy Rapin, et al. Code llama: Open foundation models for code. *arXiv preprint arXiv:2308.12950*, 2023.

Ata Çelen, Guo Han, Konrad Schindler, Luc Van Gool, Iro Armeni, Anton Obukhov, and Xi Wang. I-design: Personalized llm interior designer. *arXiv preprint arXiv:2404.02838*, 2024.

David J Tritton. Experiments on the flow past a circular cylinder at low reynolds numbers. *Journal of Fluid Mechanics*, 6(4):547–567, 1959.

Ali Kashefi and Anne E Staples. A finite-element coarse-grid projection method for incompressible flow simulations. *Advances in Computational Mathematics*, 44:1063–1090, 2018.

Ali Kashefi. A coarse grid projection method for accelerating free and forced convection heat transfer computations. *Results in Mathematics*, 75:1–24, 2020a.

Ali Kashefi. A coarse-grid projection method for accelerating incompressible mhd flow simulations. *Engineering with Computers*, 38(Suppl 2):1539–1553, 2022.

Milton Van Dyke and Milton Van Dyke. *An album of fluid motion*, volume 176. Parabolic Press Stanford, 1982.

John Kim and Parviz Moin. Application of a fractional-step method to incompressible navier-stokes equations. *Journal of computational physics*, 59(2):308–323, 1985.

Bassem F Armaly, F Durst, JCF Pereira, and B Schönung. Experimental and theoretical investigation of backward-facing step flow. *Journal of fluid Mechanics*, 127:473–496, 1983.

Hung Le, Parviz Moin, and John Kim. Direct numerical simulation of turbulent flow over a backward-facing step. *Journal of fluid mechanics*, 330:349–374, 1997.

Ali Kashefi. Coarse grid projection methodology: a partial mesh refinement tool for incompressible flow simulations. *Bulletin of the Iranian Mathematical Society*, 46:177–181, 2020b.

Leonhard Euler. Principes généraux du mouvement des fluides. *Mémoires de l'académie des sciences de Berlin*, pages 274–315, 1757.

Grzegorz Łukaszewicz and Piotr Kalita. *Navier–Stokes Equations: An Introduction with Applications*. Springer, 2016.

Peter Constantin and Ciprian Foiaş. *Navier-stokes equations*. University of Chicago press, 1988.

Alexandre Joel Chorin, Jerrold E Marsden, and Jerrold E Marsden. *A mathematical introduction to fluid mechanics*, volume 3. Springer, 1990.

George Keith Batchelor. *An introduction to fluid dynamics*. Cambridge university press, 1967.

Miloslav Feistauer, Jiří Felcman, and Ivan Straškraba. *Mathematical and computational methods for compressible flow*. Oxford University Press, USA, 2003.

Antonin Novotny and Ivan Straskraba. *Introduction to the mathematical theory of compressible flow*, volume 27. OUP Oxford, 2004.

Andrew J Majda, Andrea L Bertozzi, and A Ogawa. Vorticity and incompressible flow. cambridge texts in applied mathematics. *Appl. Mech. Rev.*, 55(4):B77–B78, 2002.

Ronald L Panton. *Incompressible flow*. John Wiley & Sons, 2024.

John E Matsson. *An Introduction to Ansys Fluent 2023*. Sdc Publications, 2023.

Amy Henderson Squillacote, James Ahrens, Charles Law, Berk Geveci, Kenneth Moreland, and Brad King. *The paraview guide*, volume 366. Kitware Clifton Park, NY, 2007.

Utkarsh Ayachit. *The paraview guide: a parallel visualization application*. Kitware, Inc., 2015.

Won Y Yang, Wenwu Cao, Jaekwon Kim, Kyung W Park, Ho-Hyun Park, Jingon Joung, Jong-Suk Ro, Han L Lee, Cheol-Ho Hong, and Taeho Im. *Applied numerical methods using MATLAB*. John Wiley & Sons, 2020.

Hrvoje Jasak. Openfoam: Open source cfd in research and industry. *International Journal of Naval Architecture and Ocean Engineering*, 1(2):89–94, 2009.

Hrvoje Jasak, Aleksandar Jemcov, Zeljko Tukovic, et al. Openfoam: A c++ library for complex physics simulations. In *International workshop on coupled methods in numerical dynamics*, volume 1000, pages 1–20, 2007.

O Heidbach, M Ziegler, and D Stromeyer. Manual of the tecplot 360 add-on geostress v2. 0. 2020.

Roger W Pryor. *Multiphysics modeling using COMSOL®: a first principles approach*. Jones & Bartlett Publishers, 2009.

Steffen Flaischlen and Gregor D Wehinger. Synthetic packed-bed generation for cfd simulations: Blender vs. star-ccm+. *ChemEngineering*, 3(2):52, 2019.

Tony Mullen. *Mastering blender*. John Wiley & Sons, 2011.

H Huang, D Dabiri, and Morteza Gharib. On errors of digital particle image velocimetry. *Measurement Science and Technology*, 8(12):1427, 1997.

Jerry Westerweel, D Dabiri, and Morteza Gharib. The effect of a discrete window offset on the accuracy of cross-correlation analysis of digital piv recordings. *Experiments in fluids*, 23(1):20–28, 1997.

Christian E Willert and Morteza Gharib. Digital particle image velocimetry. *Experiments in fluids*, 10(4):181–193, 1991.

Ronald J Adrian. Particle-imaging techniques for experimental fluid mechanics. *Annual review of fluid mechanics*, 23 (1):261–304, 1991.

Cameron Tropea, Alexander L Yarin, John F Foss, et al. *Springer handbook of experimental fluid mechanics*, volume 1. Springer, 2007.

Markus Raffel, Christian E Willert, Fulvio Scarano, Christian J Kähler, Steve T Wereley, and Jürgen Kompenhans. *Particle image velocimetry: a practical guide*. springer, 2018.

Charles Hirsch. *Numerical computation of internal and external flows: The fundamentals of computational fluid dynamics*. Elsevier, 2007.

Suhas Patankar. *Numerical heat transfer and fluid flow*. CRC press, 2018.

Stanley Osher and Ronald Fedkiw. Level set methods and dynamic implicit surfaces. *Appl. Mech. Rev.*, 57(3):B15–B15, 2004.

John David Anderson and John Wendt. *Computational fluid dynamics*, volume 206. Springer, 1995.

Junuthula Narasimha Reddy and David K Gartling. *The finite element method in heat transfer and fluid dynamics*. CRC press, 2010.

Izhar Hussain Khan, Puja Sunil, Soumarup Bhattacharyya, Rahul Yadav, Kamal Poddar, and Sanjay Kumar. Flow past two rotationally oscillating cylinders. *Journal of Fluid Mechanics*, 969:A16, 2023.

Wilson Lu, Daniah Aljubaili, Tony Zahtila, Leon Chan, and Andrew Ooi. Asymmetric wakes in flows past circular cylinders confined in channels. *Journal of Fluid Mechanics*, 958:A8, 2023.

Gregory A Lawrence and Laurence Armi. Stationary internal hydraulic jumps. *Journal of Fluid Mechanics*, 936:A25, 2022.

SF Lewin and CP Caulfield. Stratified turbulent mixing in oscillating shear flows. *Journal of Fluid Mechanics*, 944:R3, 2022.

Feng-Yuan Zuo, Antonio Memmolo, and Sergio Pirozzoli. On wall pressure fluctuations in conical shock wave/turbulent boundary layer interaction. *Journal of Fluid Mechanics*, 967:A3, 2023.

Prasoon Suchandra and Devesh Ranjan. Dynamics of multilayer rayleigh–taylor instability at moderately high atwood numbers. *Journal of Fluid Mechanics*, 974:A35, 2023.

Luke FL Alventosa, Radu Cimpeanu, and Daniel M Harris. Inertio-capillary rebound of a droplet impacting a fluid bath. *Journal of Fluid Mechanics*, 958:A24, 2023.

Jacob M Turner and Jae Wook Kim. Quadrupole noise generated from a low-speed aerofoil in near-and full-stall conditions. *Journal of Fluid Mechanics*, 936:A34, 2022.

Ian Eames and Tristan Robinson. Free-surface channel flow around a square cylinder. *Journal of Fluid Mechanics*, 980: A16, 2024.

Marco De Paoli, Sergio Pirozzoli, Francesco Zonta, and Alfredo Soldati. Strong rayleigh–darcy convection regime in three-dimensional porous media. *Journal of Fluid Mechanics*, 943:A51, 2022.

Adrián Lozano-Durán and H Jane Bae. Machine learning building-block-flow wall model for large-eddy simulation. *Journal of Fluid Mechanics*, 963:A35, 2023.

Ali Borji. Generated faces in the wild: Quantitative comparison of stable diffusion, midjourney and dall-e 2. *arXiv preprint arXiv:2210.00586*, 2022.

Dena Magdy Hanna. The use of artificial intelligence art generator "midjourney" in artistic and advertising creativity. *Journal of Design Sciences and Applied Arts*, 4(2):42–58, 2023.

Jiayang Huang, Yiran Chen, and David Yip. Crossing of the dream fantasy: Ai technique application for visualizing a fictional character's dream. In *2023 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, pages 338–342. IEEE, 2023.

Peter Henderson, Xuechen Li, Dan Jurafsky, Tatsunori Hashimoto, Mark A Lemley, and Percy Liang. Foundation models and fair use. *arXiv preprint arXiv:2303.15715*, 2023.

Soha Mansour. Intelligent graphic design: The effectiveness of midjourney as a participant in a creative brainstorming session. *International Design Journal*, 13(5):501–512, 2023.

Anna Jaruga-Rozdolska. Artificial intelligence as part of future practices in the architect's work: Midjourney generative tool as part of a process of creating an architectural form. *Architectus*, (3 (71), 2022.

Hu Yin, Zipeng Zhang, and Yuanyuan Liu. The exploration of integrating the midjourney artificial intelligence generated content tool into design systems to direct designers towards future-oriented innovation. *Systems*, 11(12):566, 2023.

Joohoon Kang and Youngjoo Yi. Beyond chatgpt: Multimodal generative ai for l2 writers. *Journal of Second Language Writing*, 62:101070, 2023.

Evelina Leivada, Elliot Murphy, and Gary Marcus. Dall· e 2 fails to reliably capture common syntactic processes. *Social Sciences & Humanities Open*, 8(1):100648, 2023.

Gary Marcus, Ernest Davis, and Scott Aaronson. A very preliminary analysis of dall-e 2. *arXiv preprint arXiv:2204.13807*, 2022.

James Betker, Gabriel Goh, Li Jing, Tim Brooks, Jianfeng Wang, Linjie Li, Long Ouyang, Juntang Zhuang, Joyce Lee, Yufei Guo, et al. Improving image generation with better captions. *Computer Science. https://cdn. openai. com/papers/dall-e-3. pdf*, 2(3):8, 2023.

Lisa C Adams, Felix Busch, Daniel Truhn, Marcus R Makowski, Hugo JWL Aerts, and Keno K Bressem. What does dall-e 2 know about radiology? *Journal of Medical Internet Research*, 25:e43110, 2023.

JaeYoung Hwang and SangHoon Oh. A brief survey of watermarks in generative ai. In *2023 14th International Conference on Information and Communication Technology Convergence (ICTC)*, pages 1157–1160. IEEE, 2023.

P Ajmera, N Nischal, S Ariyaratne, B Botchu, KDP Bhamidipaty, KP Iyengar, SR Ajmera, N Jenko, and R Botchu. Validity of chatgpt-generated musculoskeletal images. *Skeletal Radiology*, pages 1–11, 2024.

Andrew Pearson. The rise of crealtives: Using ai to enable and speed up the creative process. *Journal of AI, Robotics & Workplace Automation*, 2(2):101–114, 2023.

Chris Hales. Artificial intelligence: The latent revolution in filmmaking. *ADAM ARTS*, 2, 2021.

Imran Othman. Ai video editor: A conceptual review in generative arts. In *Proceedings of the 3rd International Conference on Creative Multimedia 2023 (ICCM 2023)*, volume 786, page 16. Springer Nature, 2023.

Sakshi Gupta. *Agency, Trust, and Interpretability of Generative Adversarial Networks (GANs)*. PhD thesis, North Carolina State University, 2021.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695, 2022.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, et al. Scaling rectified flow transformers for high-resolution image synthesis. *arXiv preprint arXiv:2403.03206*, 2024.

Chenlin Meng, Robin Rombach, Ruiqi Gao, Diederik Kingma, Stefano Ermon, Jonathan Ho, and Tim Salimans. On distillation of guided diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14297–14306, 2023.

Jonathan Ho, William Chan, Chitwan Saharia, Jay Whang, Ruiqi Gao, Alexey Gritsenko, Diederik P Kingma, Ben Poole, Mohammad Norouzi, David J Fleet, et al. Imagen video: High definition video generation with diffusion models. *arXiv preprint arXiv:2210.02303*, 2022.

Mouayad Masalkhi, Joshua Ong, Ethan Waisberg, and Andrew G Lee. Google deepmind's gemini ai versus chatgpt: a comparative analysis in ophthalmology. *Eye*, pages 1–6, 2024.

Timothy R McIntosh, Teo Susnjak, Tong Liu, Paul Watters, and Malka N Halgamuge. From google gemini to openai q*(q-star): A survey of reshaping the generative artificial intelligence (ai) research landscape. *arXiv preprint arXiv:2312.10868*, 2023.

Shashi Kant Singh, Shubham Kumar, and Pawan Singh Mehra. Chat gpt & google bard ai: A review. In *2023 International Conference on IoT, Communication and Automation Technology (ICICAT)*, pages 1–6. IEEE, 2023.

Ömer Aydin. Google bard generated literature review: metaverse. *Journal of AI*, 7(1):1–14, 2023.

Oswald Campesato. *Google Gemini for Python: Coding with Bard*. Stylus Publishing, LLC, 2024.

Matteo Mario Carlà, Gloria Gambini, Antonio Baldascino, Federico Giannuzzi, Francesco Boselli, Emanuele Crincoli, Nicola Claudio D'Onofrio, and Stanislao Rizzo. Exploring ai-chatbots' capability to suggest surgical planning in ophthalmology: Chatgpt versus google gemini analysis of retinal detachment cases. *British Journal of Ophthalmology*, 2024.

Alessio Buscemi and Daniele Proverbio. Chatgpt vs gemini vs llama on multilingual sentiment analysis. *arXiv preprint arXiv:2402.01715*, 2024.

Chaochao Lu, Chen Qian, Guodong Zheng, Hongxing Fan, Hongzhi Gao, Jie Zhang, Jing Shao, Jingyi Deng, Jinlan Fu, Kexin Huang, et al. From gpt-4 to gemini and beyond: Assessing the landscape of mllms on generalizability, trustworthiness and causality through four modalities. *arXiv preprint arXiv:2401.15071*, 2024.

Manish Bhatt, Sahana Chennabasappa, Yue Li, Cyrus Nikolaidis, Daniel Song, Shengye Wan, Faizan Ahmad, Cornelius Aschermann, Yaohui Chen, Dhaval Kapil, et al. Cyberseceval 2: A wide-ranging cybersecurity evaluation suite for large language models. *arXiv preprint arXiv:2404.13161*, 2024.

Shengrui Li, Xueting Han, and Jing Bai. Nuteprune: Efficient progressive pruning with numerous teachers for large language models. *arXiv preprint arXiv:2402.09773*, 2024a.

Jeremy Harper. Autogenesisagent: Self-generating multi-agent systems for complex tasks. *arXiv preprint arXiv:2404.17017*, 2024.

Lena Ivannova Ruiz-Rojas, Patricia Acosta-Vargas, Javier De-Moreta-Llovet, and Mario Gonzalez-Rodriguez. Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. *Sustainability*, 15(15):11524, 2023.

Hsi Yeh Wang and Surapong Utama. Investigating the generative-ai evaluation methods and correlation with fashion designers. In *2023 7th International Conference on Information Technology (InCIT)*, pages 508–513. IEEE, 2023.

Daniel Moreno-Sánchez, Daniel Moreno-Nieto, Pedro Burgos-Pintos, and Sergio I Molina. Artificial intelligence in the design workflow. review of tools with special focus on additive manufacturing. In *International conference on The Digital Transformation in the Graphic Engineering*, pages 468–475. Springer, 2023.

Nikolaus Rott. Note on the history of the reynolds number. *Annual review of fluid mechanics*, 22(1):1–12, 1990.

TA Johnson and VC Patel. Flow past a sphere up to a reynolds number of 300. *Journal of Fluid Mechanics*, 378:19–70, 1999.

Edward M Purcell. Life at low reynolds number. In *Physics and our world: reissue of the proceedings of a symposium in honor of Victor F Weisskopf*, pages 47–67. World Scientific, 2014.

Alexander J Smits, Beverley J McKeon, and Ivan Marusic. High–reynolds number wall turbulence. *Annual Review of Fluid Mechanics*, 43:353–375, 2011.

Mark Drela. Xfoil: An analysis and design system for low reynolds number airfoils. In *Low Reynolds Number Aerodynamics: Proceedings of the Conference Notre Dame, Indiana, USA, 5–7 June 1989*, pages 1–12. Springer, 1989.

Ali Kashefi, Davis Rempe, and Leonidas J Guibas. A point-cloud deep learning framework for prediction of fluid flow fields on irregular geometries. *Physics of Fluids*, 33(2), 2021.

Ali Kashefi. A coarse-grid incremental pressure projection method for accelerating low reynolds number incompressible flow simulations. *Iran Journal of Computer Science*, 3(1):13–23, 2020c.

Theodore Von Kármán. *Aerodynamics: selected topics in the light of their historical development*. Courier Corporation, 2004.

Atsushi Okajima. Strouhal numbers of rectangular cylinders. *Journal of Fluid mechanics*, 123:379–398, 1982.

WG Cochran. The flow due to a rotating disc. In *Mathematical proceedings of the Cambridge philosophical society*, volume 30, pages 365–375. Cambridge University Press, 1934.

Th V Kármán. Über laminare und turbulente reibung. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 1(4):233–252, 1921.

B.R. Munson, T.H. Okiishi, W.W. Huebsch, and A.P. Rothmayer. *Fundamentals of Fluid Mechanics*. Wiley, 2013. ISBN 9781118116135. URL `https://books.google.com/books?id=R17cthBdckYC`.

Hans G Hornung, Christian Willert, and Stewart Turner. The flow field downstream of a hydraulic jump. *Journal of Fluid Mechanics*, 287:299–316, 1995.

Gerd Baumgarten and David C Fritts. Quantifying kelvin-helmholtz instability dynamics observed in noctilucent clouds: 1. methods and observations. *Journal of Geophysical Research: Atmospheres*, 119(15):9324–9337, 2014.

RT Davis and WJ Chyu. Laminar flow past a sphere at high mach number. *Journal of Fluid Mechanics*, 24(3):481–495, 1966.

MS Howe. Attenuation of sound in a low mach number nozzle flow. *Journal of Fluid Mechanics*, 91(2):209–229, 1979.

Tengfei Luo, Jianchun Wang, Chenyue Xie, Minping Wan, and Shiyi Chen. Effects of compressibility and atwood number on the single-mode rayleigh-taylor instability. *Physics of fluids*, 32(1), 2020.

Robert M Ker. Simulation of rayleigh-taylor flows using vortex blobs. *Journal of Computational Physics*, 76(1):48–84, 1988.

Malcolm J Andrews and Stuart B Dalziel. Small atwood number rayleigh–taylor experiments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1916):1663–1679, 2010.

Ira H Abbott, Albert E Von Doenhoff, and Louis Stivers Jr. Summary of airfoil data. Technical report, 1945.

David A Burzynski, Ilia V Roisman, and Stephan E Bansmer. On the splashing of high-speed drops impacting a dry surface. *Journal of Fluid Mechanics*, 892:A2, 2020.

Haotian Liu, Chunyuan Li, Yuheng Li, and Yong Jae Lee. Improved baselines with visual instruction tuning. *arXiv preprint arXiv:2310.03744*, 2023a.

Bin Lin, Bin Zhu, Yang Ye, Munan Ning, Peng Jin, and Li Yuan. Video-llava: Learning united visual representation by alignment before projection. *arXiv preprint arXiv:2311.10122*, 2023.

C Li, C Wong, S Zhang, et al. Llava-med: training a large language-and-vision assistant for biomedicine in one day. arxiv. *arXiv preprint arXiv:2306.00890*, 2023.

Shilong Liu, Hao Cheng, Haotian Liu, Hao Zhang, Feng Li, Tianhe Ren, Xueyan Zou, Jianwei Yang, Hang Su, Jun Zhu, et al. Llava-plus: Learning to use tools for creating multimodal agents. *arXiv preprint arXiv:2311.05437*, 2023b.

Jiahui Gao, Renjie Pi, Jipeng Zhang, Jiacheng Ye, Wanjun Zhong, Yufei Wang, Lanqing Hong, Jianhua Han, Hang Xu, Zhenguo Li, et al. G-llava: Solving geometric problem with multi-modal large language model. *arXiv preprint arXiv:2312.11370*, 2023.

Jie Ren, Han Xu, Pengfei He, Yingqian Cui, Shenglai Zeng, Jiankun Zhang, Hongzhi Wen, Jiayuan Ding, Hui Liu, Yi Chang, et al. Copyright protection in generative ai: A technical perspective. *arXiv preprint arXiv:2402.02333*, 2024.

Deyao Zhu, Jun Chen, Kilichbek Haydarov, Xiaoqian Shen, Wenxuan Zhang, and Mohamed Elhoseiny. Chatgpt asks, blip-2 answers: Automatic questioning towards enriched visual descriptions. *arXiv preprint arXiv:2303.06594*, 2023.

OV Johnson, O Mohammed Alyasiri, D Akhtom, and OE Johnson. Image analysis through the lens of chatgpt-4. *Journal of Applied Artificial Intelligence*, 4(2), 2023.

James N Steer, O Kimmoun, and F Dias. Breaking-wave induced pressure and acceleration on a clifftop boulder. *Journal of Fluid Mechanics*, 929:R1, 2021.

Paul A Brandner, James A Venning, and Bryce W Pearce. Nucleation effects on cavitation about a sphere. *Journal of Fluid Mechanics*, 946, 2022.

Hang Zhang, Xin Li, and Lidong Bing. Video-llama: An instruction-tuned audio-visual language model for video understanding. *arXiv preprint arXiv:2306.02858*, 2023.

Bertalan Meskó and Eric J Topol. The imperative for regulatory oversight of large language models (or generative ai) in healthcare. *NPJ digital medicine*, 6(1):120, 2023.

Fadel M Megahed, Ying-Ju Chen, Joshua A Ferris, Sven Knoth, and L Allison Jones-Farmer. How generative ai models such as chatgpt can be (mis) used in spc practice, education, and research? an exploratory study. *Quality Engineering*, 36(2):287–315, 2024.

Lucas Theis, Aäron van den Oord, and Matthias Bethge. A note on the evaluation of generative models. *arXiv preprint arXiv:1511.01844*, 2015.

Li-Chia Yang and Alexander Lerch. On the evaluation of generative models in music. *Neural Computing and Applications*, 32(9):4773–4784, 2020.

Tianju Xue, Thomas J Wallin, Yigit Menguc, Sigrid Adriaenssens, and Maurizio Chiaramonte. Machine learning generative models for automatic design of multi-material 3d printed composite solids. *Extreme Mechanics Letters*, 41:100992, 2020.

Lyle Regenwetter, Amin Heyrani Nobari, and Faez Ahmed. Deep generative models in engineering design: A review. *Journal of Mechanical Design*, 144(7):071704, 2022.

Ali Kashefi and Tapan Mukerji. Physics-informed pointnet: A deep learning solver for steady-state incompressible flows and thermal fields on multiple sets of irregular geometries. *Journal of Computational Physics*, 468:111510, 2022.

Ali Kashefi, Leonidas J Guibas, and Tapan Mukerji. Physics-informed pointnet: On how many irregular geometries can it solve an inverse problem simultaneously? application to linear elasticity. *Journal of Machine Learning for Modeling and Computing*, 4(4), 2023.

Soumen Pal, Manojit Bhattacharya, Sang-Soo Lee, and Chiranjib Chakraborty. A domain-specific next-generation large language model (llm) or chatgpt is required for biomedical engineering and research. *Annals of Biomedical Engineering*, 52(3):451–454, 2024.

Bo Ni and Markus J Buehler. Mechagents: Large language model multi-agent collaborations can solve mechanics problems, generate new data, and integrate knowledge. *Extreme Mechanics Letters*, page 102131, 2024.

Giulia Polverini and Bor Gregorcic. How understanding large language models can inform the use of chatgpt in physics education. *European Journal of Physics*, 45(2):025701, 2024.

Rachel K Luu and Markus J Buehler. Bioinspiredllm: Conversational large language model for the mechanics of biological and bio-inspired materials. *Advanced Science*, 11(10):2306724, 2024.

Bai Li, Xinyuan Li, Yaodong Cui, Xuepeng Bian, Siyu Teng, Siji Ma, Lili Fan, Yonglin Tian, Fei-Yue Wang, et al. Integrating large language models and metaverse in autonomous racing: An education-oriented perspective. *IEEE Transactions on Intelligent Vehicles*, 2024b.