Reproducibility Review of: An Efficient System for Automatic Map Storytelling – A Case Study on Historical Maps

Item	Value
Title	An Efficient System for Automatic Map Storytelling – A Case Study on Historical Maps
Authors	Ziyi Liu, Claudio Affolter, Sidi Wu, Yizi Chen, Lorenz Hurni
Ref. paper	Liu, Z., Affolter, C., Wu, S., Chen, Y., and Hurni, L.: An Efficient System for Automatic Map Storytelling: A Case Study on Historical Maps, AGILE GIScience Ser., 6, 5, https://doi.org/10.5194/agile-giss-6-5-2025, 2025.
Codechecker(s)	Sophie Teichmann
Date of Check	2025 -04-02
Summary	Partial reproduction
Repository	https://github.com/STeichmann/automatic-map-storytelling
Ref. certificate	DOI 10.17605/OSF.IO/GT5BW

Table 1: Reproduction metadata

Output	Comment
Table 2 recreated Caption category | Base CLIP | Fine-tuned CLIP Map type | 0.43 | 0.96 (original model on Ubuntu) Location (topo) | 0.29 | 0.15 (retrained on Ubuntu) /0.85(retrained on Windows) /0.78 (orginal model on Ubuntu) Style | % | % Century | 0.41 | 0.68 (retrained on Ubuntu) /0.76 (original model) Location(pict) | 0.96 | 0.91 (original model on Ubuntu) Topic | % | % Average accuracy | % | %	Recreated Table 2 Scripts used: Inference.py
Table 3 recreated Test map | Caption category | Base CLIP | Fine-tuned  Map type | pictorial map | pictorial map Location (pict.) | world | world topic | world war 2 | flight network  Map Type | pictorial map | topographic map Location (topo.) | eastern hemisphere | asia Style | hand colored with decorative elements and pictorial relief | hand colored Century | 18th century | 19th century	Recreated Table 3 Scripts used: Inference.py (BaseCLIP) CaptionInferenceLight (Fine-tuned)
Table 4 recreated Test map   Ours (GPT-4o) | This pictorial map portrays a global flight network, detailing air routes worldwide. Set in the mid-20th century, it illustrates the interconnectedness of international air travel, serving as a tool for travelers and aviation enthusiasts to visualize and plan travel routes across continents. | This 19th-century hand-colored topographic map of Asia provides a detailed depiction of the continent's physical features during that era, offering valuable insights for historical studies, educational purposes, and geographical analysis. ClipClap | % | %	Recreated Table 4 Used demo version provided by authors

Test map		
Ours (GPT-3.5-turbo)	A 19th-century topographic map of France, hand-colored. This detailed map depicts the physical features and elevation of the land in France. It could be used for military planning, geographical research, or understanding the terrain for various purposes.	A hand-colored topographic map of France from the 19th century. It provides detailed elevation and terrain information, valuable for navigation, agriculture, urban planning, and military strategy in historical France.
Ours (GPT-4o)	This hand-colored topographic map, depicting France in the 19th century, details the country's diverse geographical features and elevations. It serves as an invaluable tool for historical researchers, educators, and enthusiasts interested in exploring France's topographical developments during this period.	This hand-colored topographic map from the 19th century depicts France, showcasing its natural features and elevations. It could be used for studying geographical features, planning, and navigation during that historical period.
Ours (GPT-4o-mini)	This hand-colored topographic map of France from the 19th century illustrates the country's diverse landscapes and elevations. It could be used for studies in geography, urban planning, and land use, providing insights into the topography of that era.	This hand-colored topographic map of France from the 19th century illustrates the varied landscapes and elevations of the region. It could be used for geographical study, historical research, and understanding land use during that period in French history.
GPT-4-turbo	This historical map depicts "Gallia Vetus," representing ancient Gaul during the Augustan era with its Roman administrative divisions. Centered on present-day France and its surrounding regions, the map details territories, primary road networks, and key cities, serving educational and historical research purposes.	This historical map illustrates the geography and principal railway networks of Italy, likely from the late 19th to early 20th century. It serves as a tool for understanding historical transit routes and regional developments within Italy during this period
GPT-4o	This map depicts ancient Gaul, corresponding to modern-day France, highlighting territorial divisions and Roman influence. It dates to the period of Roman occupation, around 58 BC. The map is useful for historical studies of Roman expansion and ancient geography.	This historical map depicts France and its surrounding regions in the 19th century, illustrating political boundaries and geographic features. It serves as a reference for historical research or educational purposes, highlighting administrative divisions and terrain during that period.
GPT-4o-mini	This historical map represents France during the early 19th century, illustrating administrative divisions and geographic features. It serves as a resource for understanding territorial organization and historical geography in France, useful for educators, researchers, and enthusiasts of French history.	This map depicts historical administrative divisions of Romania around the year 1850. It serves as a reference for understanding the geopolitical landscape of the region and can be utilized for educational, research, and historical analysis purposes.

Recreated Table 5
Used demo version provided by authors

Table 6 recreated

Test map		
Ours (GPT-3.5-turbo)	A 19th-century topographic map of Greece, hand colored, displays detailed terrain features. The map covers the geographical region of Greece, providing valuable information for military planning, land surveying, and geographical studies during the 19th century	A 19th-century hand-colored topographic map of Greece. It depicts the topography of Greece, color-coded by elevation, in the 1800s. This map could be essential for strategizing military campaigns, planning infrastructure development, or understanding the geography of Greece during that period.
Ours (GPT-4o)	This hand-colored topographic map of 19th century Greece details the region's geographical features, showcasing elevation and terrain intricacies. It serves as a vital tool for historical research, exploration, and education regarding Greece's landscape during that period.	This 19th-century hand-colored topographic map of Greece details its terrain, offering insight into the geographical landscape during that period and serving as a valuable resource for historical research, education, and understanding the country's topography and development in that era.
Ours (GPT-4o-mini)	This 19th-century topographic map of Greece is hand-colored, showcasing the region's varied terrain and geographic features. It serves as a valuable resource for historians, geographers, and travelers interested in understanding Greece's topography during that period.	This 19th-century topographic map of Greece is hand-colored, detailing the country's varied landscapes. It serves to illustrate geographical features and can be utilized for educational purposes, historical research, or aiding travelers in understanding Greece's terrain during that period.
GPT-4-turbo	This map depicts Ancient Greece, highlighting its geographical regions and principal cities in a historical context. Useful for educational purposes, it provides a detailed representation of Greece's terrain and settlements during antiquity, facilitating deeper understanding of its historical and cultural development.	This map depicts Italy as it was around the year 1200, showcasing regional divisions, topography, and key cities. Designed for historical and educational purposes, it offers insights into Italy's medieval geopolitical landscape which can be useful for researchers and educators delving into historical geography.
GPT-4o	This map depicts Ancient Greece, highlighting geographical regions and city-states, and includes an inset of the vicinity of Athens. It is a historical map useful for studying the geography and political boundaries of ancient Greece, likely from the classical period.	This map depicts Italy as it was around the year 1200, showcasing regional divisions, topography, and key cities. Designed for historical and educational purposes, it offers insights into Italy's medieval geopolitical landscape which can be useful for researchers and educators delving into historical geography.
GPT-4o-mini	This map depicts ancient Greece, highlighting its geographic features and settlements. It is focused on the region of Greece and its surrounding islands, likely from the 19th century. The map serves educational purposes, aiding in historical	This map depicts the region of Italy as it was around the year 1200, illustrating its geographical features and political boundaries. It can be utilized for historical studies and understanding medieval

Recreated Table 6
Used demo version provided by authors.

Table 7 recreated

Test map		
Ours (GPT-3.5-turbo)	A pictorial map of the world during World War 2. This map showcases global locations and events during the war in the early to mid-20th century, offering a visual reference for historical study and understanding of the conflict's impact globally.	A pictorial map depicting the world during World War 2. The map showcases countries involved in the conflict, major battle sites, and key events. This historical map serves as a visual aid to study the global impact of World War 2.
Ours (GPT-4o)	This pictorial map depicts the global dynamics during World War 2, illustrating various locations worldwide between 1939 and 1945. It serves as an educational tool for understanding geopolitical changes, military campaigns, and strategic alliances of that era, enhancing visual comprehension of historical events.	This pictorial map depicts the world during World War 2, illustrating significant geopolitical changes and events. It provides an educational tool to understand wartime dynamics and strategic developments globally.
Ours (GPT-4o-mini)	The pictorial map of the world during World War II creatively illustrates the global landscape, highlighting significant battles, alliances, and nations involved. This map serves as an educational tool, providing historical context and visual representation of wartime dynamics from 1939 to 1945.	The pictorial map of the world illustrates the global landscape during World War II, detailing key events and locations involved in the conflict. This map serves as a historical reference for understanding the geopolitical dynamics of the era.
GPT-4-turbo	This map, titled "Bacon's Standard Map of the World on Mercators Projection," is a global depiction highlighting Allied and Axis territories during World War II, showcasing control as of April 1942. It serves as a historical reference to visualize geopolitical affiliations and military progress during the war.	This map depicts the Cold War era, focusing on the global division of power and influence between the Western and Eastern blocs. It highlights territories aligned with NATO, the Warsaw Pact, and non-aligned nations, providing a visual representation of political affiliations and military alliances during the mid-to-late 20th century. This map could be employed for educational contexts to explain geopolitical dynamics and tensions of the period.
GPT-4o	This map, titled "Hammond's War Map," depicts global political boundaries and alliances around 1942 during World War II. It highlights Axis and Allied powers, neutral countries, and colonies. It was likely used for educational or informational purposes during the war to illustrate global geopolitical dynamics.	This map depicts the global colonial empires circa 1910, illustrating the territorial holdings of major powers across continents. It serves educational purposes, highlighting the political geography of imperialism and colonial influence during the early 20th century, particularly useful for historical analyses and studies.
GPT-4o-mini	This map, titled "Baker's 1943 War Map," depicts global military conflicts during World War II, focusing on various nations and regions. It serves as a historical reference for understanding wartime alliances and territorial control.	This map illustrates global political divisions and territorial claims during the early 20th century, specifically around 1940. It serves as a historical reference for understanding geopolitical contexts and colonial influences across various regions.

Recreated Table 7
Used demo version provided by the authors

Table 2: Summary of output files generated

Summary

The main challenge during this reproduction is the combination of a closed source model (Chat GPT) and available models (like CLIP). Using the models provided by the authors, the results were reproducible (Recreated Table 2). Using Ubuntu the results were partly not recreatable on Ubuntu. On Windows the training was successful. Thus, there is a system dependence of the reproducibility when retraining the models.

Another challenge was the volume of models and corresponding inference. Thus I only chose a subset of models to retrain and perform inference on. A part of the inference in the paper was done manually in the paper, which I was not able to recreate.

I was able to fully recreate Table1, Table 3, Table 5, Table 6 and Table 7 from the original paper. I would count the overall reproduction as a success.

Reproducibility reviewer notes

The GitHub repository provided by the authors (https://github.com/claudaff/automatic-map-storytelling) is well structured and documented and provides good guidance. They also give access to a free demo version upon request (for the combined GPT models with their fine-tuned models). The authors provided the maps used in the paper (augemented and non-augemented).

I was not able to attempt the reproduction of training all models, I can only make a limited statement about the reproducibility of the training. For the three scripts I tried, the training was completed both on Ubuntu and also on Windows. The fine-tuning for the locations of the topographic maps was only successful on Windows.

Overall when using the models the authors provided, the models performed as expected. As expected the captions generated with large language models were not identical, however they where very similar and don't seem to contradict the authors conclusions.

Installation prerequisites and computational environment

I was able to recreate the environment according to the GitHub repository of the authors with python 3.9.12 on Ubuntu 24.04.2. Switched during the review to Windows 11 and python 3.9.13. Most of the review was done on Ubuntu, if Windows was used it is explicitly mentioned. The authors provided an easy set-up with Anaconda. For me, the set-up failed (due to installation of the required version scikit-image) for python 3.12.3 (Ubuntu) and 3.11.5 (Windows).

Data preparation

The authors used data from the online map repository David Rumsey Map Collection (https://www.davidrumsey.com/). The data is a map with a corresponding captions file.

They provide their selection of training data as a download. Pre-processing of the data is done by the scripts “CaptionGenerationPictorial.py” and “CaptionGenerationClassical.py”.

The resulting array with the data (map type, map data, short title, path, etc), was successfully created. This was validated because I was able to recreate the data in Table 1 by printing the length of the variables the quantity corresponds to.

Running the code

I was not able to recreate all of the training due to the amount computational resources and time constraints. The scripts look well structured and well documented. The authors provide six scripts for the respective training of the fine-tuned models. The authors provide the fine-tuned models in the their repository. I ran the script for fine-tuning to century (run time: 1:20:34) and style (run time: 1:15:47) and location (topo) (run time (ubunutu: 57:10, Windows: 1:19:27). The best model on Ubuntu is reached after a few epochs and after that the cross entropy loss of the validation set does not improve anymore. It might be reasonable to employ early stopping after the validation loss does not decrease further after ~5 epochs to reduce computation time.

Because I had no available openAI API key, I was also not able to test “CaptionInferenceGUI.py”.

I was only able to recreate a few of the resulting figures due to the mentioned computational restraints.

Table 2 orginal

Table 2. Comparison of prediction accuracies achieved per caption category with the base CLIP model and our fine-tuned CLIP models.

Caption category	Base CLIP	Fine-tuned CLIP
Map type	0.43	**0.96**
Location (topo.)	0.28	**0.78**
Style	0.29	**0.75**
Century	0.40	**0.76**
Location (pict.)	**0.96**	0.93
Topic	0.47	**0.67**
Average Accuracy	0.47	**0.81**

Table 2 recreated

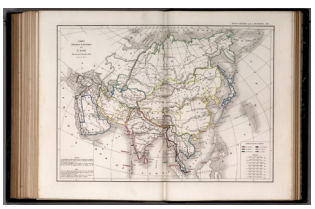
Caption category	Base CLIP	Fine-tuned CLIP
Map type	0.43	0.96 (original model on Ubuntu)
Location (topo)	0.29	0.15 (retrained on Ubuntu) /0.85(retrained on Windows) /0.78 (orginal model on Ubuntu)
Style	%	%
Century	0.41	0.68 (retrained on Ubuntu) /0.76 (original model)
Location(pict)	0.96	0.91 (original model on Ubuntu)
Topic	%	%
Average accuracy	%	%

The results were recreated using the “Inference.py” script with uncommenting the corresponding lines. Since the provided path for the test set were windows paths, I added a line to convert the paths to the format needed in Ubuntu. Since no ground truth was provided for the categories topic and style, I was not able to calculate an accuracy. The authors did the evaluation manually. There is difference between the model trained on Ubuntu and the models in the paper. The authors provided the information that this is likely due to misalignment of the labels created in the “CaptionGenerationClassic.py” script so the fine tuning is done on mislabeled data and thus not effective. On Windows, I was able to get an even higher accuracy when performing the retraining.

Table 3 original

Test map	Caption category	Base CLIP	Fine-tuned
	Map type	pictorial map	pictorial map
	Location (pict.)	world	world
	Topic	world war 2	flight network
	Map type	pictorial map	topographic map
	Location (topo.)	eastern hemisphere	asia
	Style	hand colored with decorative elements and pictorial relief	hand colored
	Century	18th century	19th century

Table 3 recreated

Test map	Caption category	Base CLIP	Fine-tuned
	Map type	pictorial map	pictorial map
	Location (pict.)	world	world
	topic	world war 2	flight network
	Map Type	pictorial map	topographic map
	Location (topo.)	eastern hemisphere	asia
	Style	hand colored with decorative elements and pictorial relief	hand colored
	Century	18th century	19th century

The authors provided a the script “CaptionInferenceLight.py” that recreated the outputs of the Fine-Tuned model. The base CLIP inference was made with the script ‘Inference.py’ on Ubuntu setting ‘useFineTuned = False’ and uncommenting the right labels for the caption category.

Table 4 original

Table 4. The stories of the same test maps generated by our method and ClipCap (Mokady et al., 2021). ⁕: Fine-tuned ClipCap.

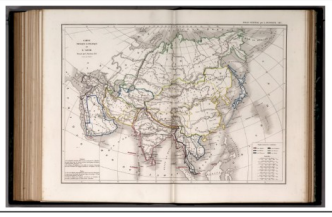
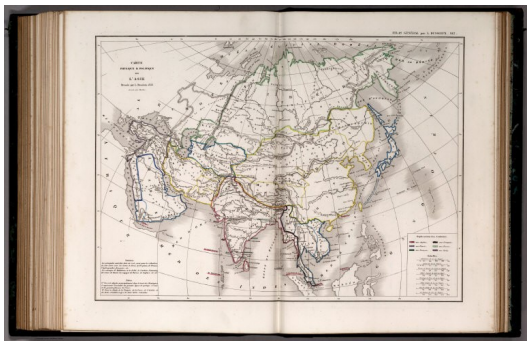
Test map		
Ours	This pictorial map illustrates the global flight network, showcasing worldwide destinations and travel routes. It is a visual representation of the world, providing information about flight connections, and can be used for planning and visualizing travel itineraries.	This hand-colored topographic map of Europe in the 19th century features pictorial relief. It shows the geographical features of Europe and can be used for geographical analysis.
ClipCap	A map of the world is on display in a room.	An old map with a map cutter on it.
ClipCap⁕	Map depicting Air France worldwide flight network.	Map depicting Europe in the 17th century.

Table 4 recreated

Test map		
Ours (GPT-4o)	This pictorial map portrays a global flight network, detailing air routes worldwide. Set in the mid-20th century, it illustrates the interconnectedness of international air travel, serving as a tool for travelers and aviation enthusiasts to visualize and plan travel routes across continents.	This 19th-century hand-colored topographic map of Asia provides a detailed depiction of the continent's physical features during that era, offering valuable insights for historical studies, educational purposes, and geographical analysis.
ClipClap	%	%

It is expected to gain similar but not identical outputs for the corresponding maps. However, a major difference is the but the map is correctly identified as Asian in our

approach but not in original paper. Unfortunately, the interface to use the ClipClap (https://huggingface.co/spaces/akhaliq/CLIP_prefix_captioning) model had technical issues during the review so no comparison was possible.

Table 5 original

Table 5. Comparison of generated captions between our system integrating different GPT versions for storytelling and the vision-enabled GPTs that can generate captions directly, when texts on maps are missing. On the right, the title and production year (at bottom-left) are covered. Falsely predicted information is marked in red.

Test map		
Ours (GPT-3.5-turbo)	A hand-colored topographic map of France from the 19th century showcases the terrain and features for geographic analysis, military planning, and navigation purposes within the country during that era.	A 19th-century hand-colored topographic map of France. The map depicts the geographical features and terrain of France during that time and could be used for navigation, military planning, and understanding the country's landscape in the 19th century.
Ours (GPT-4o)	This hand-colored topographic map of France from the 19th century provides detailed elevation and landscape features of the region, offering insights into geographic terrains applicable to studies in historical geography, exploration, and cartography during that period.	This topographic map, hand-colored and depicting France in the 19th century, provides detailed geographical and elevation information, ideal for understanding the landscape and planning developments or travel strategies of that era across the nation's diverse regions.
Ours (GPT-4o-mini)	This hand-colored topographic map represents France in the 19th century. It portrays the country's geographical features, including elevation and terrain. Such maps could be used for historical research, education, and understanding of regional development during that period.	This 19th-century hand-colored topographic map illustrates the geographical features of France. It serves as a valuable resource for understanding the country's terrain, aiding in studies related to geography, history, and land use during that period.
GPT-4-turbo	This historical map, titled "Gallia Vetus," represents Ancient Gaul during Roman times, detailing its diverse tribal regions and main settlements. It serves educational and research purposes, offering insights into ancient geographic and political boundaries and the development of early European civilizations.	This historical map illustrates the railway networks across Spain and Portugal as of 1900. It details rail lines, geographical features, and cities, making it a valuable tool for understanding regional connectivity and travel planning during that era.
GPT-4o	This map depicts Ancient Gaul, focusing on geographical and political regions in what is now France, during antiquity. It serves as a historical reference to understand the territorial organization and tribal distributions of that era.	This historical map depicts France and surrounding regions, focusing on political boundaries and geographical features of the time, around the 19th century. It serves as a tool for understanding historical territorial divisions and geographical knowledge during this period.
GPT-4o-mini	This map illustrates the region of Gaul as it appeared in 1820. It serves to depict administrative divisions, geographic features, and transportation routes, useful for historians, geographers, and those studying early 19th-century France.	This map depicts Romania, showcasing its administrative divisions and geographical features. It is from the 19th century and can be used for historical research and understanding regional developments during that time.

Table 5 recreated

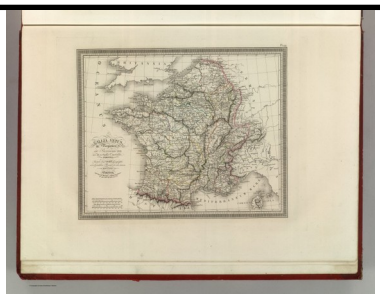
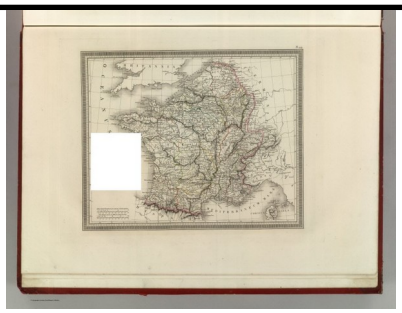
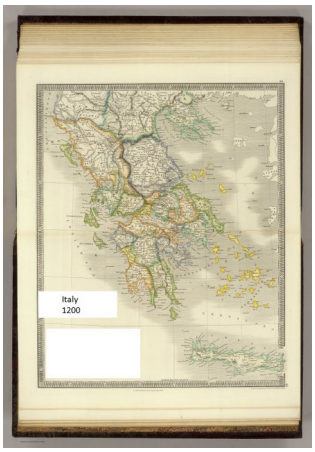
Test map		
Ours (GPT-3.5-turbo)	A 19th-century topographic map of France, hand-colored. This detailed map depicts the physical features and elevation of the land in France. It could be used for military planning, geographical research, or understanding the terrain for various purposes.	A hand-colored topographic map of France from the 19th century. It provides detailed elevation and terrain information, valuable for navigation, agriculture, urban planning, and military strategy in historical France.
Ours (GPT-4o)	This hand-colored topographic map, depicting France in the 19th century, details the country's diverse geographical features and elevations. It serves as an invaluable tool for historical researchers, educators, and enthusiasts interested in exploring France's topographical developments during this period.	This hand-colored topographic map from the 19th century depicts France, showcasing its natural features and elevations. It could be used for studying geographical features, planning, and navigation during that historical period.
Ours (GPT-4o-mini)	This hand-colored topographic map of France from the 19th century illustrates the country's diverse landscapes and elevations. It could be used for studies in geography, urban planning, and land use, providing insights into the topography of that era.	This hand-colored topographic map of France from the 19th century illustrates the varied landscapes and elevations of the region. It could be used for geographical study, historical research, and understanding land use during that period in French history.
GPT-4-turbo	This historical map depicts "Gallia Vetus," representing ancient Gaul during the Augustan era with its Roman administrative divisions. Centered on present-day France and its surrounding regions, the map details territories, primary road networks, and key cities, serving educational and historical research purposes.	This historical map illustrates the geography and principal railway networks of Italy, likely from the late 19th to early 20th century. It serves as a tool for understanding historical transit routes and regional developments within Italy during this period
GPT-4o	This map depicts ancient Gaul, corresponding to modern-day France, highlighting territorial divisions and Roman influence. It dates to the period of Roman occupation, around 58 BC. The map is useful for historical studies of Roman expansion and ancient geography.	This historical map depicts France and its surrounding regions in the 19th century, illustrating political boundaries and geographic features. It serves as a reference for historical research or educational purposes, highlighting administrative divisions and terrain during that period.
GPT-4o-mini	This historical map represents France during the early 19th century, illustrating administrative divisions and geographic features. It serves as a resource for understanding territorial organization and historical geography in France, useful for educators, researchers, and enthusiasts of French history.	This map depicts historical administrative divisions of Romania around the year 1850. It serves as a reference for understanding the geopolitical landscape of the region and can be utilized for educational, research, and historical analysis purposes.

Table 6 original

Table 6. Comparison of generated captions between our system integrating different GPT versions for storytelling and the vision-enabled GPTs that can generate captions directly, when texts on maps are incorrect. On the right, wrong title and production year are injected. Falsely predicted information is marked in red.

Test map		
Ours (GPT-3.5-turbo)	A 19th-century hand-colored topographic map of Greece could be utilized for geographical analysis, land surveying, and historical research, providing detailed visual information about the terrain features, elevations, and geographic characteristics of Greece as depicted in the 1800s.	A hand-colored topographic map of Greece from the 19th century depicts detailed physical features. It provides information on the country's terrain, elevations, and natural landscape, aiding in military planning, infrastructure development, and geographical studies of Greece during that era.
Ours (GPT-4o)	This hand-colored topographic map depicts the geographical features of Greece during the 19th century. It provides detailed elevation and terrain information, which could be used for historical research, educational purposes, or understanding the landscape and development of Greece in that era.	The hand-colored topographic map of 19th-century Greece illustrates the country's geographical features and elevations. Created during this period, it could be utilized for military planning, educational purposes, or enhancing understanding of Greece's diverse landscape in that era.
Ours (GPT-4o-mini)	This 19th-century topographic map of Greece is hand-colored, depicting the region's geographic features. It can be used for educational purposes, historical reference, or research into the landscape and topography of Greece during that era.	This 19th-century topographic map of Greece is hand-colored, showcasing the region's varied landscape. It can be used for geographical studies, historical research, or to understand the terrain and settlements of Greece during that period.
GPT-4-turbo	This map illustrates Ancient Greece along with surrounding island groups and inset maps of the Athenian Empire and the Environs of Athens. This historical map, rendered in detail, could serve educational and academic purposes, aiding in the study of Greek history and geography.	This map represents Italy in the year 1200, detailing its geographical features and political divisions during the medieval period. It serves as an educational tool for understanding Italy's historical territorial boundaries and regional distinctions at that time.
GPT-4o	This map depicts Ancient Greece, focusing on its geographical regions and city-states. It provides a historical overview, likely from the classical period, and can be used for educational purposes or historical research on ancient Greek territories and their locations.	The map depicts Greece around the year 1200, highlighting regional divisions and geographic features. This historical map can be used for educational purposes, illustrating political boundaries and geography during the period, aiding in studies of medieval Greek history and cartography.
GPT-4o-mini	This map depicts Ancient Greece, focusing on its geographic features and city-states. The map illustrates the region's landscape during classical antiquity, serving as a historical reference for educational and research purposes.	This map depicts Greece in the year 1200. It illustrates territorial divisions and geographic features, valuable for historical research or education regarding medieval Greek topography and settlement patterns.

Table 6 recreated

Test map		
Ours (GPT-3.5-turbo)	A 19th-century topographic map of Greece, hand colored, displays detailed terrain features. The map covers the geographical region of Greece, providing valuable information for military planning, land surveying, and geographical studies during the 19th century	A 19th-century hand-colored topographic map of Greece. It depicts the topography of Greece, color-coded by elevation, in the 1800s. This map could be essential for strategizing military campaigns, planning infrastructure development, or understanding the geography of Greece during that period.
Ours (GPT-4o)	This hand-colored topographic map of 19th century Greece details the region's geographical features, showcasing elevation and terrain intricacies. It serves as a vital tool for historical research, exploration, and education regarding Greece's landscape during that period.	This 19th-century hand-colored topographic map of Greece details its terrain, offering insight into the geographical landscape during that period and serving as a valuable resource for historical research, education, and understanding the country's topography and development in that era.
Ours (GPT-4o-mini)	This 19th-century topographic map of Greece is hand-colored, showcasing the region's varied terrain and geographic features. It serves as a valuable resource for historians, geographers, and travelers interested in understanding Greece's topography during that period.	This 19th-century topographic map of Greece is hand-colored, detailing the country's varied landscapes. It serves to illustrate geographical features and can be utilized for educational purposes, historical research, or aiding travelers in understanding Greece's terrain during that period.
GPT-4-turbo	This map depicts Ancient Greece, highlighting its geographical regions and principal cities in a historical context. Useful for educational purposes, it provides a detailed representation of Greece's terrain and settlements during antiquity, facilitating deeper understanding of its historical and cultural development.	This map depicts Italy as it was around the year 1200, showcasing regional divisions, topography, and key cities. Designed for historical and educational purposes, it offers insights into Italy's medieval geopolitical landscape which can be useful for researchers and educators delving into historical geography.
GPT-4o	This map depicts Ancient Greece, highlighting geographical regions and city-states, and includes an inset of the vicinity of Athens. It is a historical map useful for studying the geography and political boundaries of ancient Greece, likely from the classical period.	This map depicts Italy as it was around the year 1200, showcasing regional divisions, topography, and key cities. Designed for historical and educational purposes, it offers insights into Italy's medieval geopolitical landscape which can be useful for researchers and educators delving into historical geography.
GPT-4o-mini	This map depicts ancient Greece, highlighting its geographic features and settlements. It is focused on the region of Greece and its surrounding islands, likely from the 19th century. The map serves educational purposes, aiding in historical	This map depicts the region of Italy as it was around the year 1200, illustrating its geographical features and political boundaries. It can be utilized for historical studies and understanding medieval

	studies and understanding of ancient Greek civilization.	Italian territorial organization.

Table 7 orginal

Table 7. Comparison of generated captions between our system integrating different GPT versions for storytelling and the vision-enabled GPTs that can generate captions directly, when texts on maps are altered. On the right, the title and map legend are erased using an AI tool. Falsely predicted information is marked in red.

Test map		
Ours (GPT-3.5-turbo)	A pictorial map of the world during World War 2. The map illustrates the global theater of the conflict, showcasing key locations, battles, and events from the war. It serves as a visual aid for understanding the scope and impact of World War 2.	A pictorial map of the world during World War 2, depicting key locations and events. It could be used for educational purposes or as a historical reference to visualize the global impact of the war.
Ours (GPT-4o)	This pictorial map of the world during World War 2 provides a visual representation of global events and territorial boundaries. It illustrates the geopolitical landscape of the era and could be used for educational purposes, understanding historical context, or visualizing wartime strategies.	This pictorial map illustrates global strategic locations and events of World War 2. It depicts the world during the period of World War 2 and can be used for educational purposes to understand the geopolitical dynamics and military campaigns of that era.
Ours (GPT-4o-mini)	The pictorial map of the world during World War II illustrates the geopolitical landscape and significant events of the time. It focuses on global movements and conflicts from 1939 to 1945, serving as an educational tool for understanding wartime dynamics.	This pictorial map of the world illustrates key events and themes during World War II. It focuses on various countries affected by the war, serving as a visual reference for historical education and analysis of global conflict from 1939 to 1945.
GPT-4-turbo	This map, titled "Baird North's War Map," provides a comprehensive view of World War II status across the globe. It visually delineates territorial control and alliances as of 1942. It serves educational and historical analysis uses, detailing key regions, battles, and occupation zones worldwide.	This map depicts the Cold War era, highlighting the global political divisions between the Western Bloc, the Eastern Bloc, and Non-Aligned countries. It covers the entire world and is useful for understanding mid-20th-century geopolitical alignments and tensions.
GPT-4o	This map depicts global territorial control during World War II as of 1942. It highlights Axis, Allied, and neutral countries. Created for educational purposes, it provides a visual representation of geopolitical power dynamics and wartime alliances during this period.	This 1939 world map illustrates British Empire territories in red, showcasing colonial dominance across continents. It serves educational or political purposes, highlighting the global extent of British control at that time, reflecting geopolitical conditions shortly before World War II.
GPT-4o-mini	This map illustrates global military theaters during World War II, specifically from 1942. It serves to depict zones of conflict and military presence, aiding in geographical understanding of wartime strategies and alliances across various continents.	This map illustrates historical colonial empires and their territories around the world, focusing on the mid-20th century. It serves to highlight global colonial influences and could be used for educational purposes regarding imperial history and geographical changes.

Table 7 recreated

Test map		
Ours (GPT-3.5-turbo)	A pictorial map of the world during World War 2. This map showcases global locations and events during the war in the early to mid-20th century, offering a visual reference for historical study and understanding of the conflict's impact globally.	A pictorial map depicting the world during World War 2. The map showcases countries involved in the conflict, major battle sites, and key events. This historical map serves as a visual aid to study the global impact of World War 2.
Ours (GPT-4o)	This pictorial map depicts the global dynamics during World War 2, illustrating various locations worldwide between 1939 and 1945. It serves as an educational tool for understanding geopolitical changes, military campaigns, and strategic alliances of that era, enhancing visual comprehension of historical events.	This pictorial map depicts the world during World War 2, illustrating significant geopolitical changes and events. It provides an educational tool to understand wartime dynamics and strategic developments globally.
Ours (GPT-4o-mini)	The pictorial map of the world during World War II creatively illustrates the global landscape, highlighting significant battles, alliances, and nations involved. This map serves as an educational tool, providing historical context and visual representation of wartime dynamics from 1939 to 1945.	The pictorial map of the world illustrates the global landscape during World War II, detailing key events and locations involved in the conflict. This map serves as a historical reference for understanding the geopolitical dynamics of the era.
GPT-4-turbo	This map, titled "Bacon's Standard Map of the World on Mercators Projection," is a global depiction highlighting Allied and Axis territories during World War II, showcasing control as of April 1942. It serves as a historical reference to visualize geopolitical affiliations and military progress during the war.	This map depicts the Cold War era, focusing on the global division of power and influence between the Western and Eastern blocs. It highlights territories aligned with NATO, the Warsaw Pact, and non-aligned nations, providing a visual representation of political affiliations and military alliances during the mid-to-late 20th century. This map could be employed for educational contexts to explain geopolitical dynamics and tensions of the period.
GPT-4o	This map, titled "Hammond's War Map," depicts global political boundaries and alliances around 1942 during World War II. It highlights Axis and Allied powers, neutral countries, and colonies. It was likely used for educational or informational purposes during the war to illustrate global geopolitical dynamics.	This map depicts the global colonial empires circa 1910, illustrating the territorial holdings of major powers across continents. It serves educational purposes, highlighting the political geography of imperialism and colonial influence during the early 20th century, particularly useful for historical analyses and studies.
GPT-4o-mini	This map, titled "Baker's 1943 War Map," depicts global military conflicts during World War II, focusing on various nations and regions. It serves as a historical reference for understanding wartime alliances and territorial control.	This map illustrates global political divisions and territorial claims during the early 20th century, specifically around 1940. It serves as a historical reference for understanding geopolitical contexts and colonial influences across various regions.

Recommendations

Even though openAI provides state of the art models, I would suggest to switch to an open available. Furthermore, is the use of Anaconda also under license. I would recommend that the authors mention in their repository, that the generation of training labels only works reliably in Windows. It would be very helpful for future reproduction when the authors would provide a file with a list what file produces which figure/table. Additionally it would be helpful for future reproductions if the authors would provide the (augmented) figures that were used in the paper in the GitHub repository.

Acknowledgments

I want to thank the authors for the very friendly and timely communication and providing the demo version for their combined models and the interesting review. Additionally, I want to thank the AGILE reproducibility review committee for this opportunity to perform this review.

Citing this document

This report is part of the reproducibility review at the AGILE conference. For more information see https://reproducible-agile.github.io/. This document is published on OSF/RE at OSF/RE DOI 10.17605/OSF.IO/GT5BW.

To cite the report use

DOI 10.17605/OSF.IO/GT5BW