Computation and Performance Studies
Volume 1 · Number 1 · 2026

Sixteen Times Fewer Pixel Coordinates, the Same Recognition Count: A Matched-Resolution Gallery Study

Owen Carter[1] · Liam Johnston[2] · Evan Anderson[3]

[1] Poland · [2] Hungary · [3] Croatia

Original publication: October 28, 2026 | Article 671

Abstract

Reducing a CMU image from 128 by 120 pixels to its supplied 32 by 30 version left all 155 query correctness indicators unchanged in a simple gallery search. Both grids identified 151 queries correctly, while the coordinate budget fell sixteenfold. This finite agreement is not an equivalence guarantee. A subsequent exploratory analysis found that the mean separation between the closest wrong and correct identities decreased by 0.00497 per-pixel squared-intensity units, with a working interval from -0.00596 to -0.00398. Resolution reduction preserved the observed decision count while changing the geometry supporting those decisions.

Source-provided image scales are compared on identical captures. Accuracy is descriptive because every paired query correctness difference is zero; the distance-margin analysis was added afterward and is exploratory.

Sixteen Times Fewer Pixel Coordinates, the Same Recognition Count: A Matched-Resolution Gallery Study

> Source-provided image scales are compared on identical captures. Accuracy is descriptive because every paired query correctness difference is zero; the distance-margin analysis was added afterward and is exploratory.

Abstract

Reducing a CMU image from 128 by 120 pixels to its supplied 32 by 30 version left all 155 query correctness indicators unchanged in a simple gallery search. Both grids identified 151 queries correctly, while the coordinate budget fell sixteenfold. This finite agreement is not an equivalence guarantee. A subsequent exploratory analysis found that the mean separation between the closest wrong and correct identities decreased by 0.00497 per-pixel squared-intensity units, with a working interval from -0.00596 to -0.00398. Resolution reduction preserved the observed decision count while changing the geometry supporting those decisions.

A cheaper representation may preserve the count but change the evidence

Image resolution affects both the size of a numerical representation and the details available for matching. A practitioner using a simple reference gallery might therefore ask whether the largest source image is necessary for the observed task. The answer should distinguish a smaller input representation, a measured speed change, and the quality of the resulting decisions. These quantities are related but are not identical resource or performance measures.

Our primary question is what recognition change accompanies replacing native CMU captures with the corresponding source-provided quarter-size images in the same gallery search. The gallery members, query captures, identity labels, and decision rule remain fixed. The comparison examines one source-defined resolution policy rather than an optimized family of resizing filters or a learned compression method.

The first result was unusually simple: every query's correct-or-incorrect status remained unchanged. We therefore retain accuracy as an observed finite-cohort result and do not fit a degenerate mean-difference model with zero standard error. After seeing that outcome, we added a separate exploratory question about distance margins. Its timing and status are explicit so that a nonzero result is not presented as a prespecified replacement for the unchanged primary endpoint.

Meng et al., 2026 use stacked multi-view identity mosaics for subject-preserving video generation. Reference representations can create resource questions in that setting, but this manuscript does not implement a mosaic, encode a video, or generate a subject. It studies a simpler measurement tradeoff: whether a smaller whole-image representation preserves roster lookup decisions and the numerical separation underlying them.

Three meanings of a low-dimensional image

Downsampling a source image is different from projecting it onto a learned representation. The early closest-fit perspective in Pearson, 1901 and probabilistic principal components in Tipping and Bishop, 1999 describe forms of representation through fitted directions. Turk and Pentland, 1991 and Belhumeur et al., 1997 apply learned image representations to recognition. No projection matrix is estimated in our comparison; we read two provided pixel grids directly.

Modern identity embeddings provide another route. Schroff et al., 2015 learns a metric, while Kim et al., 2022 and Meng et al., 2021 relate representation learning to image quality. Their output dimension is not simply the number of input pixels, and their computation includes a learned forward pass. A sixteenfold reduction in our raw pixel coordinates therefore says nothing by itself about the memory or latency of those networks.

Detection also precedes many real recognition pipelines. Deng et al., 2019 concerns face localization, whereas our source images enter the distance calculation as complete rectangular frames. A detector could change both the useful image content and the resource balance. Its absence keeps the experiment reproducible but prevents the results from being interpreted as an end-to-end production face-recognition benchmark.

The representation distinction extends to generative video. Blattmann et al., 2023 studies scaling latent video diffusion to large datasets, where resolution, latent structure, and output quality have different computational implications. The present study cannot infer those tradeoffs from a small grayscale gallery. Its relevance is the need to name the representation and measured operation before associating fewer coordinates with a quality-preserving efficiency claim.

Paired captures at two supplied sizes

The data source is Mitchell, 1997, with original educational documentation in Carnegie Mellon University, 2026. The archive provides 624 usable captures of twenty people after sixteen source-marked camera glitches are excluded. Each usable capture has native, half-size, and quarter-size PGM variants. We decode their actual dimensions and retain a capture identifier linking the variants, rather than treating each size as a new photograph.

Only open-eye images enter this analysis. Neutral and happy expressions form the gallery across available head poses, yielding 158 references. Sad and angry expressions form 155 queries over the same twenty identities. The few missing source captures remain absent. The protocol neither replaces them nor selects a more favorable complete subset after computing results.

For a given resolution, every query is compared with the same roster of reference captures at that resolution. Native images have 120 rows and 128 columns, or 15,360 grayscale coordinates. Quarter-size images have thirty rows and thirty-two columns, or 960 coordinates. The reduction is sixteenfold in pixel count, although the source uses quarter to describe each spatial dimension rather than the total number of pixels.

We use the source's existing size derivatives instead of regenerating them with a newly selected interpolation filter. Their exact sampling and filtering history is not inferred from the suffix alone. Consequently, the comparison evaluates the delivered source-size policy, including its actual pixel values. It does not isolate a universal effect of pixel count while keeping every other aspect of image formation fixed.

No query capture is a gallery capture, but all query identities are known. This is closed-roster whole-image lookup. The source pictures can carry identity information through clothing, background, framing, and other persistent capture cues. High observed accuracy is therefore not an independent assessment of facial identity invariance or performance on unfamiliar people.

The same numerical rule on both grids

Images are represented as float64 grayscale values divided by 255. We compute squared Euclidean distances and assign the identity of the nearest gallery photograph. Exact ties follow lexicographic gallery order. There is no learned embedding, registration, image centering, or brightness normalization. The only experimental policy change is which supplied source grid is used.

The implementation uses array operations documented in Harris et al., 2020, with the inferential calculations relying on statistical routines in Virtanen et al., 2020 through the local analysis libraries. Distances are formed from squared norms and a matrix product. Small negative numerical values are floored at zero, consistent with the nonnegativity of exact squared distances. The numerical-analysis perspectives in Higham, 2002 and Goldberg, 1991 motivate recording arithmetic and precision rather than assuming every algebraically equivalent implementation behaves identically.

The output distance matrix has 24,490 entries at either resolution: 155 queries times 158 references. What changes is the dimension involved in each pairwise comparison. Its pair-coordinate product count is proportional to 376,166,400 at native size and 23,510,400 at quarter size. These are dimensions of the mathematical workload, not a measured count of processor instructions or a promise of proportional runtime.

We also measured a narrow in-memory operation twelve times at each size after initial predictions had been obtained. Execution order alternated across rounds. The timed operation includes distance construction and nearest-identity selection, using already loaded arrays; it excludes disk reading, image decoding, preprocessing, and any model training. All repeated predictions were checked against the initially stored labels.

Accuracy did not move, so its uncertainty was not manufactured

Both grids correctly identified 151 of 155 queries. The pooled accuracy was therefore about 97.42%, while the equal-person mean was 97.50% because source capture counts differ slightly by person. Every query's correctness indicator was identical across the two grids. The result is stronger than equal aggregate counts alone, but it does not establish that every incorrect identity label was also unchanged.

An intercept regression of the paired accuracy differences would have an outcome that is zero at every person. Its sample variance and ordinary standard error would be zero, while an R-squared diagnostic is undefined. Such output would not supply a valid certificate that future accuracy differences are exactly zero. We record the finite agreement directly and exclude that degenerate model from the inferential table.

Agreement perspectives in Bland and Altman, 1986 and Lin, 1989 help distinguish matching observed decisions from a broad claim that two procedures are interchangeable. An assessment on new captures, a changed gallery, or difficult boundary examples could produce different labels. The existing source roster is not a randomly sampled population of future recognition conditions, and no equivalence margin or power calculation was specified for it.

The stored predicted identities were not identical for every query, even though correctness was identical. At least one error therefore changed which wrong roster identity was returned without changing the binary success count. This distinction matters when errors have different consequences. A count-based recognition measure would treat those outputs alike, while a workflow that reviews the actual assigned identity might not. We did not attach identity-specific costs or infer which wrong assignment was more consequential, so the unchanged accuracy result remains a statement about its particular binary endpoint.

An explicitly later check of decision margins

After observing unchanged correctness, we computed an additional diagnostic. For each query, we find the smallest squared distance to any reference of its true identity and the smallest distance to a reference of another identity. Their difference, wrong minus true, is divided by the number of pixels at that resolution. A positive margin favors the true identity; the outcome is a geometry-based diagnostic, not a calibrated probability of correctness.

True query labels are used only to construct this evaluation margin. The nearest-gallery classifier does not receive them. The diagnostic asks how strongly the raw representation separates the nearest correct and competing identities on the supplied roster. It does not estimate confidence for an unknown person or set an operational rejection threshold.

Native mean person margin was 0.023428 per-pixel squared-intensity units. Quarter-size mean margin was 0.018457. We average query margin changes within person, giving twenty paired rows. An intercept-only OLS model estimates their mean, using HC3 covariance and t19 intervals. White, 1980 motivates the robust covariance; no explanatory controls or fixed effects enter the model.

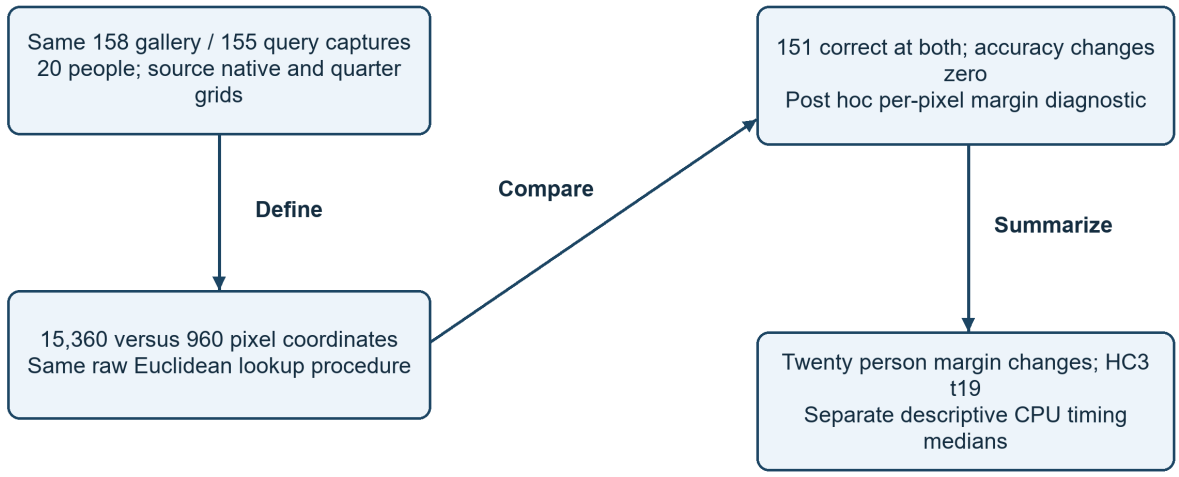


Figure 1. *Source resolution changes the distance representation while capture membership remains fixed. The margin outcome was added after the zero accuracy difference was observed; timing remains a separate descriptive measurement.*

Table 1. *Exploratory quarter-size minus native mean per-pixel identity margin. N equals twenty person means; intercept-only OLS, HC3 standard error, t19 interval, no controls or fixed effects. Accuracy differences are not included as a zero-variance inferential model.*

Model	Term	Estimate (SE)	95% CI	N
M1	Mean outcome/contrast	-0.004971 (0.000475)	[-0.005965, -0.003978]	20
M1: fit	R-squared = 0.0000	-	-	20

Model labels: M1 = Exploratory quarter-size minus native per-pixel identity margin.

The estimated margin change was -0.004971, with standard error 0.000475 and working interval -0.005965 to -0.003978. The delivered smaller grid therefore showed less separation on this particular

normalized distance diagnostic despite preserving observed correctness. A smaller margin does not itself imply lower robustness to every perturbation: no perturbation family was executed to establish such a relationship.

The inferential scope is conditional on the source people, capture menus, and post hoc endpoint. The measurement-error perspective in Carroll et al., 2006 cautions against treating a distance proxy as latent identity strength. Stevens, 1946 requires its unit to remain explicit. Per-pixel squared-intensity separation is not a percentage of identity retained or an externally validated perceptual quality score.

The smaller margin also cannot be interpreted as a measured loss of confidence. Confidence calibration would require a probability model and an assessment relating its stated probabilities to observed correctness. Here the diagnostic is the difference between two minima in a pixel metric, and both minima can change under the source resolution policy. Per-pixel normalization removes a direct dimension multiplier, but it does not make different spatial representations identical measurements of the same physical quantity.

A follow-up robustness experiment could add a fixed family of query perturbations and ask whether the two grids cross identity boundaries at different rates. That would be new evidence, with an intervention and severity scale absent from this study. We do not use the negative margin contrast as a substitute for those unobserved trials. Keeping this boundary explicit preserves the useful finding that numerical separation changed while avoiding an unsupported prediction that smaller images must fail sooner in every application.

Coordinate savings and the measured operation

Median native lookup time in this CPU session was approximately 0.011928 seconds, compared with 0.000867 seconds at quarter size. These narrow timings are descriptive; no population latency interval is attached to twelve repeated calls within one session. Results can change with thread settings, cache state, libraries, and hardware, even when the mathematical predictions remain unchanged.

Williams et al., 2009 distinguishes computational and memory constraints, explaining why dimension reduction need not translate into an identical speed ratio. Mytkowicz et al., 2009 examines measurement bias in performance experiments, and Kalibera and Jones, 2013 emphasizes replication appropriate to the performance claim. Our single-session medians are consequently not a hardware benchmark, energy comparison, or measurement of complete application response time.

The two floating-point input matrices together contain sixteen times fewer values at the smaller grid, but the distance output has the same dimensions. Program overhead and output allocation therefore do not vanish with the input coordinates. Packed image-file storage, network transmission, and low-bit representation are separate resource questions not measured by these float64 arrays.

Learned visual representations such as Radford et al., 2021, Caron et al., 2021, and Oquab et al., 2023 could change both separation and computation. Their benefits cannot be inferred by assuming a smaller raw grid is equivalent to a learned compact feature. The current result instead supplies a transparent baseline against which such a representation experiment could later be specified with actual inference and matching capture sets.

A bounded resource conclusion

System dependencies discussed in Sculley et al., 2015 matter because an apparently simple resolution choice affects input construction, numerical workload, and evaluation meaning. Documentation in Gebru et al., 2021 and reuse principles in Wilkinson et al., 2016 motivate keeping those links explicit. The local records include original capture mappings, query outcomes, exploratory margin values, timing scope, and the reason the zero-variance accuracy regression was not retained.

For this gallery and these 155 queries, the smaller supplied grid retained every correctness outcome while reducing coordinate work and narrowing the observed distance margins. That is a useful finite comparison, with no guarantee of equal future accuracy or general robustness. The appropriate engineering claim is that this specific baseline exhibited a cheaper observed representation with unchanged task counts, while its supporting geometry and unmeasured deployment costs remained distinct.

References

1. Deng, Jiankang, et al. "RetinaFace: Single-stage Dense Face Localisation in the Wild." *arXiv preprint arXiv:1905.00641*, 2019. https://arxiv.org/abs/1905.00641.
2. Lin, Lawrence I-Kuei. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics. https://doi.org/10.2307/2532051.
3. Meng, Qiang, et al. "MagFace: A Universal Representation for Face Recognition and Quality Assessment." *arXiv preprint arXiv:2103.06627*, 2021. https://arxiv.org/abs/2103.06627.
4. Pearson, Karl. (1901). LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. https://doi.org/10.1080/14786440109462720.
5. Mytkowicz, Todd; Diwan, Amer; Hauswirth, Matthias; Sweeney, Peter F.. (2009). Producing wrong data without doing anything obviously wrong!. Proceedings of the 14th international conference on Architectural support for programming languages and operating systems. https://doi.org/10.1145/1508244.1508275.
6. Turk, M.A.; Pentland, A.P.. (1991). Face recognition using eigenfaces. Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/cvpr.1991.139758.
7. Caron, Mathilde, et al. "Emerging Properties in Self-Supervised Vision Transformers" *arXiv preprint arXiv:2104.14294*, 2021. https://arxiv.org/abs/2104.14294.
8. Belhumeur, P.N.; Hespanha, J.P.; Kriegman, D.J.. (1997). Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence. https://doi.org/10.1109/34.598228.
9. Harris, Charles R.; Millman, K. Jarrod; van der Walt, Stéfan J.; Gommers, Ralf; Virtanen, Pauli; Cournapeau, David; Wieser, Eric; Taylor, Julian; Berg, Sebastian; Smith, Nathaniel J.; Kern, Robert; Picus, Matti; Hoyer, Stephan; van Kerkwijk, Marten H.; Brett, Matthew; Haldane, Allan; del Río, Jaime Fernández; Wiebe, Mark; Peterson, Pearu; Gérard-Marchant, Pierre; Sheppard, Kevin; Reddy, Tyler; Weckesser, Warren; Abbasi, Hameer; Gohlke, Christoph; Oliphant, Travis E.. (2020). Array programming with NumPy. Nature. https://doi.org/10.1038/s41586-020-2649-2.
10. Williams, Samuel; Waterman, Andrew; Patterson, David. (2009). Roofline: an insightful visual performance model for multicore architectures. Communications of the ACM. https://doi.org/10.1145/1498765.1498785.
11. Radford, Alec, et al. "Learning Transferable Visual Models From Natural Language Supervision" *arXiv preprint arXiv:2103.00020*, 2021. https://arxiv.org/abs/2103.00020.
12. Higham, Nicholas J.. (2002). Accuracy and Stability of Numerical Algorithms: Second Edition. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898718027.
13. Carnegie Mellon University. (2026). Neural Networks for Face Recognition: companion data and file documentation [Original educational archive; accessed8November2026]. https://www.cs.cmu.edu/~tom/faces.html.
14. Mitchell, Tom. (1997). CMU Face Images [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C5JC79.
15. Kim, Minchul, et al. "AdaFace: Quality Adaptive Margin for Face Recognition." *arXiv preprint arXiv:2204.00964*, 2022. https://arxiv.org/abs/2204.00964.

16. Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu. “Measurement Error in Nonlinear Models.” 2006. https://doi.org/10.1201/9781420010138.

17. Tipping, Michael E.; Bishop, Christopher M.. (1999). Probabilistic Principal Component Analysis. Journal of the Royal Statistical Society Series B: Statistical Methodology. https://doi.org/10.1111/1467-9868.00196.

18. Sculley, D., et al. "Hidden Technical Debt in Machine Learning Systems." *Advances in Neural Information Processing Systems*, vol. 28, 2015. https://papers.nips.cc/paper/2015/hash/86df7dcfd896fcaf2674f757a2463eba-Abstract.html

19. Virtanen, Pauli; Gommers, Ralf; Oliphant, Travis E.; Haberland, Matt; Reddy, Tyler; Cournapeau, David; Burovski, Evgeni; Peterson, Pearu; Weckesser, Warren; Bright, Jonathan; van der Walt, Stéfan J.; Brett, Matthew; Wilson, Joshua; Millman, K. Jarrod; Mayorov, Nikolay; Nelson, Andrew R. J.; Jones, Eric; Kern, Robert; Larson, Eric; Carey, C J; Polat, İlhan; Feng, Yu; Moore, Eric W.; VanderPlas, Jake; Laxalde, Denis; Perktold, Josef; Cimrman, Robert; Henriksen, Ian; Quintero, E. A.; Harris, Charles R.; Archibald, Anne M.; Ribeiro, Antônio H.; Pedregosa, Fabian; van Mulbregt, Paul; SciPy 1.0 Contributors; Vijaykumar, Aditya; Bardelli, Alessandro Pietro; Rothberg, Alex; Hilboll, Andreas; Kloeckner, Andreas; Scopatz, Anthony; Lee, Antony; Rokem, Ariel; Woods, C. Nathan; Fulton, Chad; Masson, Charles; Häggström, Christian; Fitzgerald, Clark; Nicholson, David A.; Hagen, David R.; Pasechnik, Dmitrii V.; Olivetti, Emanuele; Martin, Eric; Wieser, Eric; Silva, Fabrice; Lenders, Felix; Wilhelm, Florian; Young, G.; Price, Gavin A.; Ingold, Gert-Ludwig; Allen, Gregory E.; Lee, Gregory R.; Audren, Hervé; Probst, Irvin; Dietrich, Jörg P.; Silterra, Jacob; Webber, James T; Slavič, Janko; Nothman, Joel; Buchner, Johannes; Kulick, Johannes; Schönberger, Johannes L.; de Miranda Cardoso, José Vinícius; Reimer, Joscha; Harrington, Joseph; Rodríguez, Juan Luis Cano; Nunez-Iglesias, Juan; Kuczynski, Justin; Tritz, Kevin; Thoma, Martin; Newville, Matthew; Kümmerer, Matthias; Bolingbroke, Maximilian; Tartre, Michael; Pak, Mikhail; Smith, Nathaniel J.; Nowaczyk, Nikolai; Shebanov, Nikolay; Pavlyk, Oleksandr; Brodtkorb, Per A.; Lee, Perry; McGibbon, Robert T.; Feldbauer, Roman; Lewis, Sam; Tygier, Sam; Sievert, Scott; Vigna, Sebastiano; Peterson, Stefan; More, Surhud; Pudlik, Tadeusz; Oshima, Takuya; Pingel, Thomas J.; Robitaille, Thomas P.; Spura, Thomas; Jones, Thouis R.; Cera, Tim; Leslie, Tim; Zito, Tiziano; Krauss, Tom; Upadhyay, Utkarsh; Halchenko, Yaroslav O.; Vázquez-Baeza, Yoshiki. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. https://doi.org/10.1038/s41592-019-0686-2.

20. Gebru, Timnit, et al. "Datasheets for datasets" *Communications of the ACM*, 2021. https://doi.org/10.1145/3458723.

21. Schroff, Florian, et al. "FaceNet: A Unified Embedding for Face Recognition and Clustering" *arXiv preprint arXiv:1503.03832*, 2015. https://arxiv.org/abs/1503.03832.

22. Kalibera, Tomas; Jones, Richard. (2013). Rigorous benchmarking in reasonable time. Proceedings of the 2013 international symposium on memory management. https://doi.org/10.1145/2464157.2464160.

23. Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S.; Heringa, Jaap; ’t Hoen, Peter A.C; Hooft, Rob; Kuhn, Tobias; Kok, Ruben; Kok, Joost; Lusher, Scott J.; Martone, Maryann E.; Mons, Albert; Packer, Abel L.; Persson, Bengt; Rocca-Serra, Philippe; Roos, Marco; van Schaik, Rene; Sansone, Susanna-Assunta; Schultes, Erik; Sengstag, Thierry; Slater, Ted; Strawn, George; Swertz, Morris A.; Thompson, Mark; van der Lei, Johan; van Mulligen, Erik; Velterop, Jan; Waagmeester, Andra; Wittenburg, Peter; Wolstencroft, Katherine; Zhao, Jun; Mons, Barend. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. https://doi.org/10.1038/sdata.2016.18.

24. Oquab, Maxime, et al. "DINOv2: Learning Robust Visual Features without Supervision." *arXiv preprint arXiv:2304.07193*, 2023. https://arxiv.org/abs/2304.07193.

25. White, Halbert. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica*, vol. 48, no. 4, 1980, pp. 817-838. https://doi.org/10.2307/1912934.

26. Blattmann, Andreas; Dockhorn, Tim; Kulal, Sumith; Mendelevitch, Daniel; Kilian, Maciej; Lorenz, Dominik; Levi, Yam; English, Zion; Voleti, Vikram; Letts, Adam; Jampani, Varun; Rombach, Robin. (2023). Stable Video Diffusion: Scaling Latent Video Diffusion Models to Large Datasets. arXiv:2311.15127. https://arxiv.org/abs/2311.15127.

27. Goldberg, David. (1991). What every computer scientist should know about floating-point arithmetic. ACM Computing Surveys. https://doi.org/10.1145/103162.103163.

28. Stevens, S. S.. (1946). On the Theory of Scales of Measurement. Science. https://doi.org/10.1126/science.103.2684.677.

29. Bland, J. Martin, and Douglas G. Altman. "Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement." *The Lancet*, vol. 327, no. 8476, 1986, pp. 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8.

30. Meng, Zijie; Liu, Jiwen; Liu, Yufei; Tong, Chengzhuo; Liu, Xiaoqiang; Zhang, Yuanxing; Xu, Yulong; Wan, Pengfei. (2026). ARGUS: Stacked Multi-View Identity Mosaic Injection for Subject-Preserving Video Generation. arXiv:2606.11670. https://arxiv.org/abs/2606.11670.