

Creativity in Automatic Sketch-based Image Creation

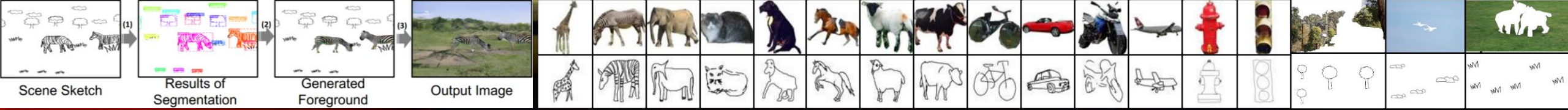
Proyecto final

Creatividad Computacional Proyecto final Yan



Introduction & Objective

- Emulation of human creativity with generative adversarial neural networks (GANs)
- SketchyGAN: Tries to generate realistic images from a sketch
- Objective: Outline how SketchyGAN could be converted from coach system into co-creative computer colleague
 - Improvisation, real-time collaboration between human & computer
 - Detailed explanation, concept for GUI
 - Techniques for colleague system could be used for other GANs too



Related Work

- No research for sketch-based image generation published by ICCV
- Most systems for category-specific sketches
 - E.g. only faces (Osahor et al., 2020), faces, birds and cars (Lu et al., 2018) or shoes and chairs (Liu et al., 2019)
- Cross-domain generalization limited (*right image*) (Liu et al., 2019)
- SketchyCOCO: multiple drawn objects (*top image*) (Gao et al., 2020)
 - 14 predefined object classes, 3 types of backgrounds
 - Segmentation, separate images, combination on background



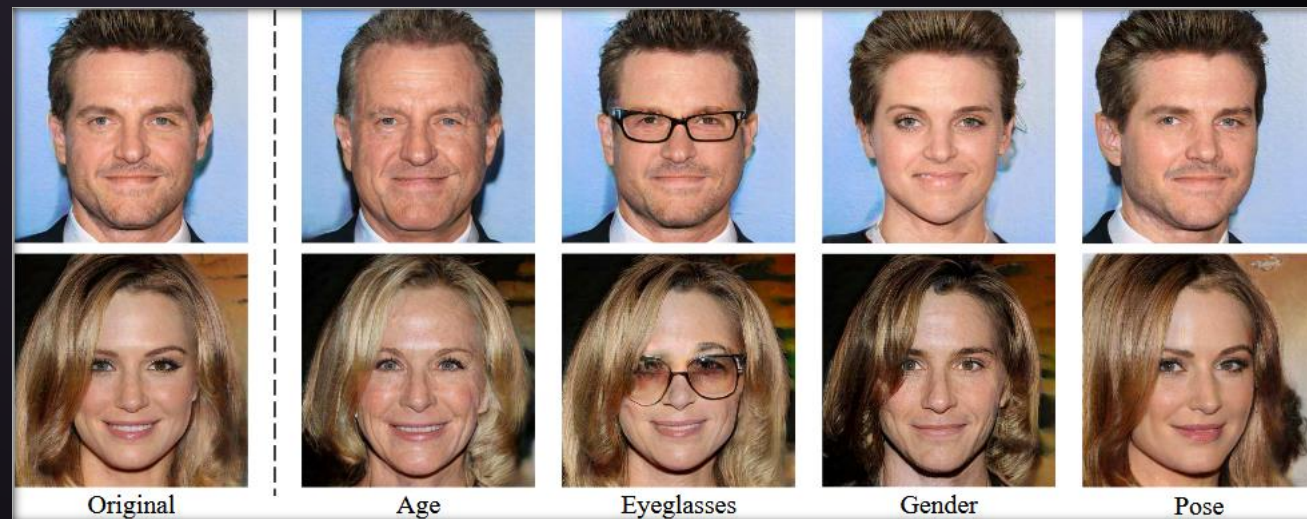
Sketch of a car
with network
trained on shoes

Related Work

- SketchyGAN (Chen & Hays, 2018) and MindReader (Guo et al., 2019) don't need specified category
 - But are trained with limited number of categories
- Human sketches are distorted and crude
 - Simple, inaccurate conceptualization of idea
 - Semantic understanding of human intent
 - Domain-specific corrections for realism
 - To augment training data, systems can distort & simplify detected edges (Liu et al., 2019)

Related Work: Creativity research

- Sketch-based GANs act as coach systems
- Demonstrate combinational creativity
 - Combine previously seen ideas of sketches & realistic representations
- Maybe exploratory creativity
 - GANs search for optimal solution
 - Latent space has semantic meaning (*image*)
(Shen et al., 2020)



SketchyGAN as autonomous creative system

- By description of Dan Ventura for computationally creative systems (Ventura, 2017)
- Domain: digital art
- Phenotype: 64x64 px for produced colored image
- Genotype: Sketch image converted to 64x64 unsigned Euclidian distance field for sketch edge map
 - Original sketch is sparse. Now each pixel has information about nearest edge, more useful data (*image*)
- Knowledge base: Training data
 - Sketchy dataset (Sangkloy et al., 2016) augmented with images from internet and their edge maps
- Aesthetic inherited from knowledge base
 - Discriminator learns aesthetic, no specific rules



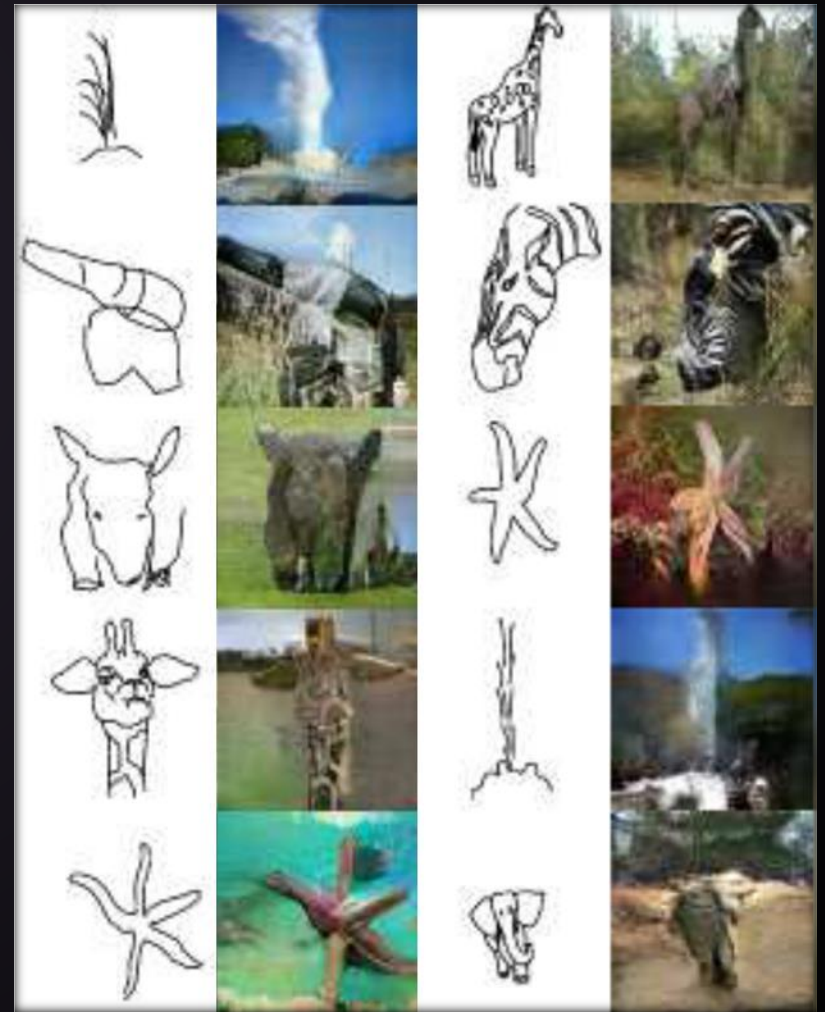
SketchyGAN as autonomous creative system

- Conceptualization: GAN has countless neurons with individual weights
- Generation with Conceptualization
 - GAN runs in cycle, generator & discriminator train each other
- Has no distinct Genotypic Evaluator
 - Aesthetic evaluation automatically learned by discriminator
- Translation: Generator function of GAN
- Phenotypic Evaluation by discriminator
 - Discriminator learns to distinguish realistic images from badly generated ones
 - Could be used to check system's products

Creative limitations of SketchyGAN

1. Creative = novel, surprising, valuable (Boden, 1998)

- Only novel to the system, that doesn't mean much
 - Concept of digital art with GAN is not new
 - No transformational creativity
- Surprising
 - A few results could be surprising to some people
 - Distortions, digital artifacts, weird results
- Valuable
 - Might be valuable to some people (digital artists)



Creative limitations of SketchyGAN

2. Definition of autonomy (Jennings, 2010)

- Autonomous evaluation fulfilled by discriminator
- Autonomous change not fulfilled
 - GAN can change the weights of its neurons, but not how it operates
 - No creative freedom to create how it wants, no self-determination
- Non-randomness mostly fulfilled
- Products are based on (preferably) human sketches => not fully autonomous

Creative limitations of SketchyGAN

3. Understanding of domain

- Intelligent, intuitive understanding or just imitation?
- Knowledge base of limited categories
 - Only understands those objects
 - Communicates certain, limited domain aesthetic
 - Can hardly transfer knowledge to other domain
 - No commonsense or knowledge of real world
- SketchyGAN often fails to understand human intent
 - *"Ideally, we want our results to be both realistic and faithful to the intent of the input sketch. For many sketches, we fail to meet one or both of these goals. Results generally aren't photorealistic..."*



Evaluation of generated products

- Would fail modified Turing test (low resolution, digital artifacts, distortions)
- Appropriate for digital art domain
- Typicality subjective, surveys needed to quantify
- Novelty also subjective (art defined by cultural interpretation and societal norms)
 - Some novelty due to unpredictable properties of some results, dissimilarity to other digital art
 - Disparity between products and human (more realistic) interpretation of the sketch
- Value highly subjective, survey with domain experts recommended
- Products generate some surprise
- Further research: Consensual assessment with domain experts and general population

Process assessment

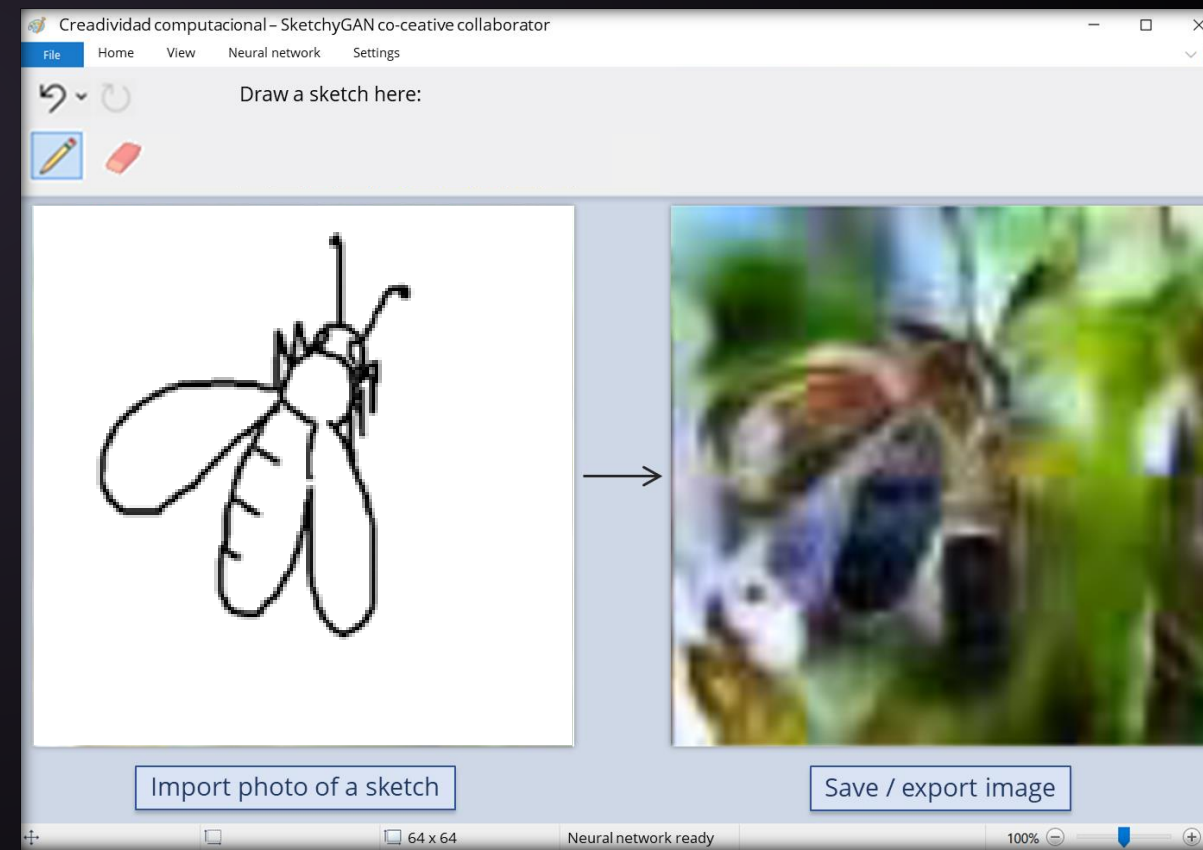
- Some randomness in how result is distorted / changed
 - Same input = same result; slightly changed input = completely different result
 - Products orient on input sketch, often faithful to intent
- Intent hard to determine
 - SketchyGAN cannot evaluate its own process, but its products
 - Generation-evaluation cycle constantly improves system
- Level of creativity: “Generalization” stage
 - Has autonomy to change its results, can create variation:
 - By design generalizes over training data, no capacity to memorize knowledge base
- Maybe “Filtration” stage, if intent can be proven



Co-creative colleague version

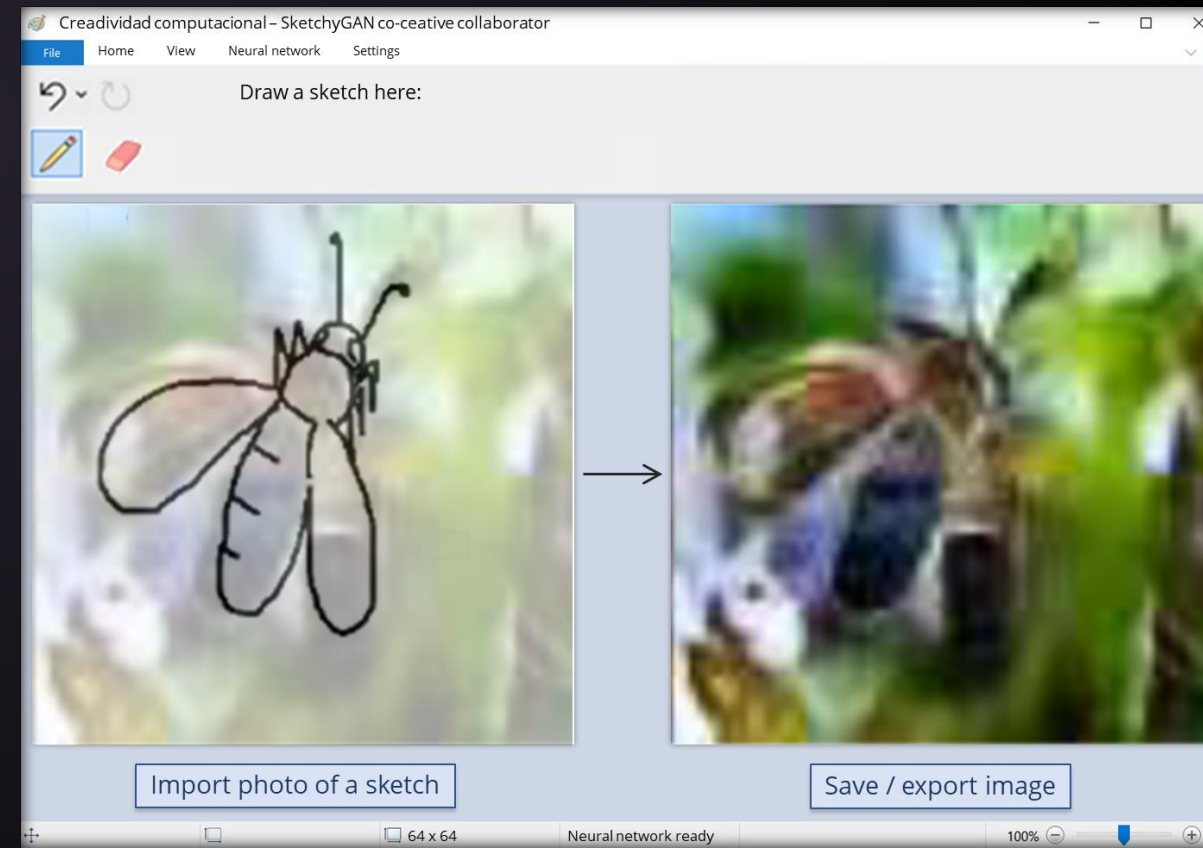
Goal: Co-creative collaboration between human & system

- Real-time sketch manipulation in canvas
 - With every change system updates product
 - Enables real-time collaboration
 - Multi-threaded implementation, so user doesn't have to wait on product update
- Undo, redo, eraser give user possibilities & creative freedom to explore ideas without repercussions



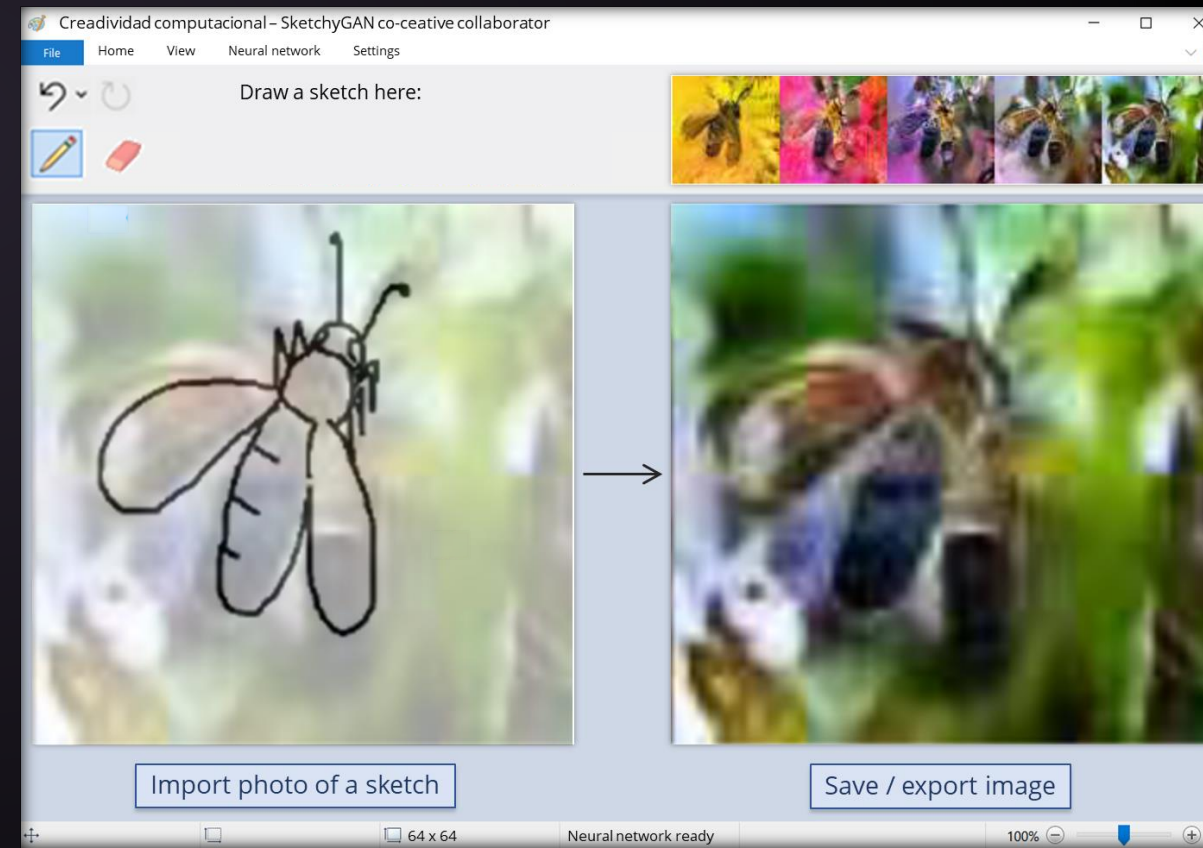
Co-creative colleague version

- Transparent sketch overlay
 - Input and output on the same canvas
 - More direct interaction with product
 - Can correct distorted lines in the sketch
 - User and system suggest each other new ideas, contribute at the same time
 - System shapes user's directive, vice versa



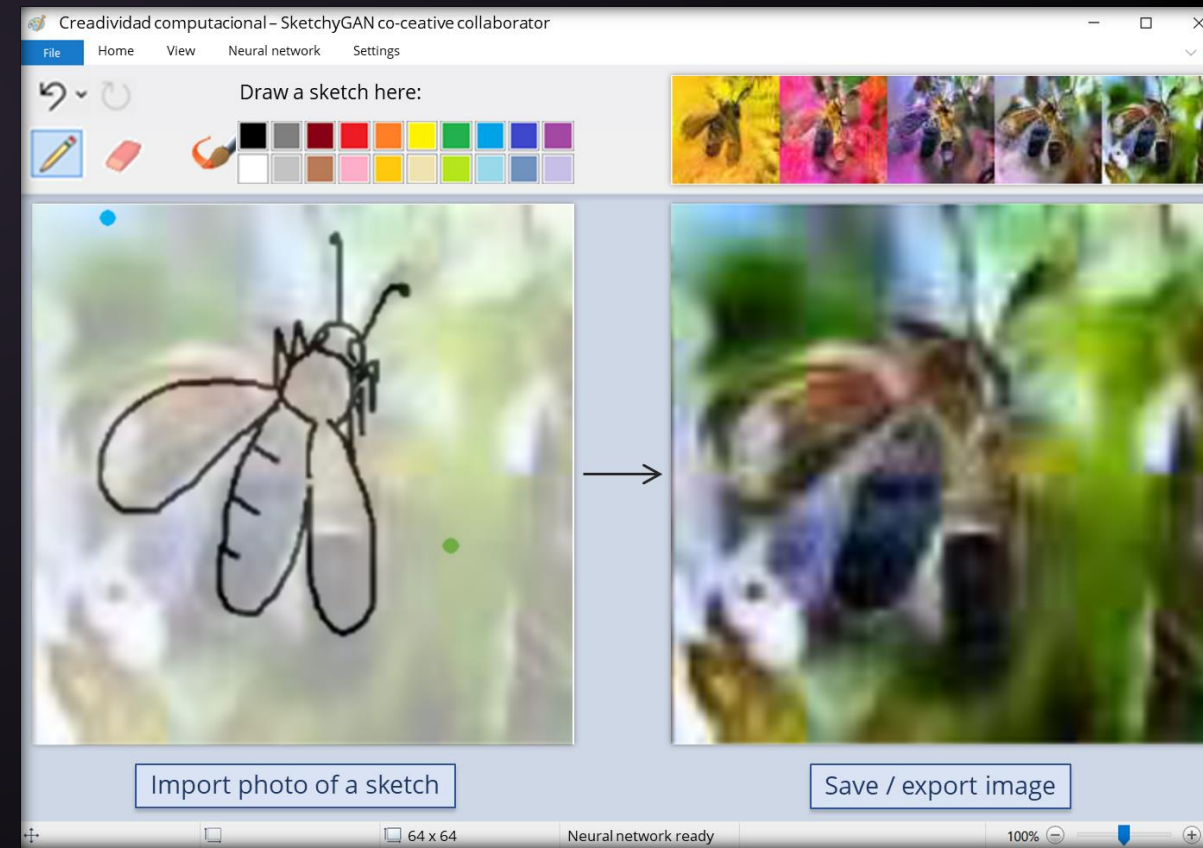
Co-creative colleague version

- Multiple product suggestions
 - For exploration of many different ideas
 - Helps shape new directives
 - Slightly different noise vectors on input can produce different products
 - User selects sketch that best fits directive
- Problem: After selecting background, sketch can't be changed again without resulting in completely new products with new backgr.



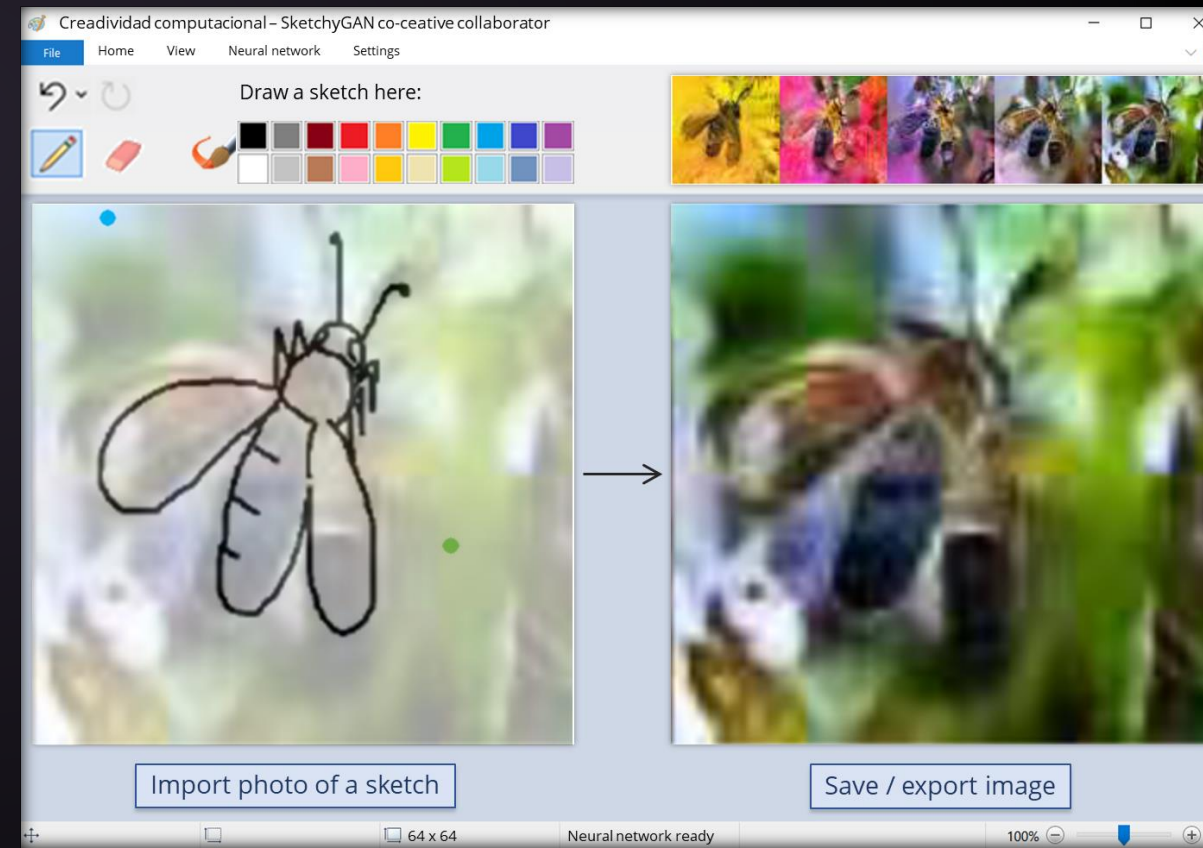
Co-creative colleague version

- Sketch doesn't contain info about background
- Solution: Filtering by color points
 - User defines color points on parts of sketch
 - Products are filtered for that preference (similar color in that area)
 - Continue editing without losing preference
 - More creative freedom for user, influence over other parts of image
 - System has more ability to make good suggestions to preferred aesthetic



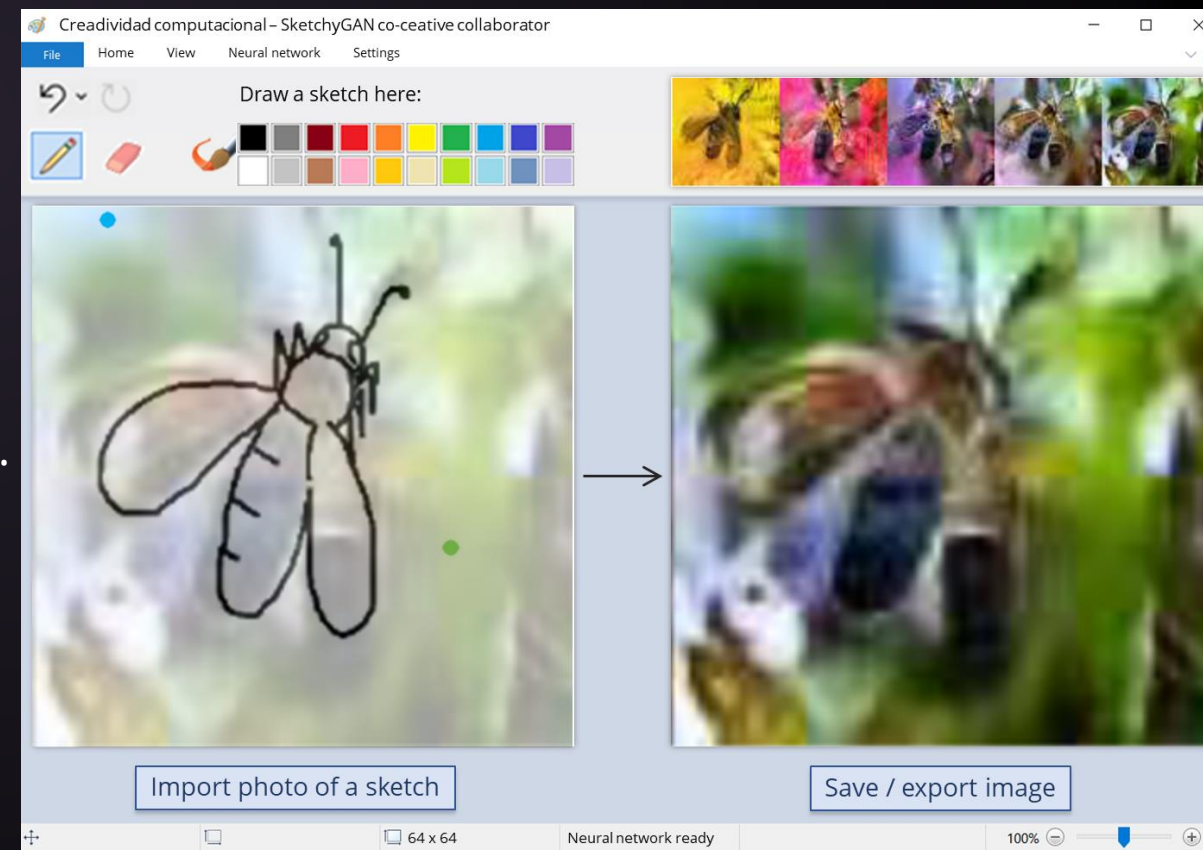
Co-creative colleague version

- With too many color points could be hard to find appropriate products
 - Genotype now includes color points
 - Train GAN with randomly chosen color points from target image
 - Fitness function will accept similar color at these locations in generated image
 - Trains GAN to better understand human intent about image features



Co-creative colleague version

- Conclusion: Human & system work together on product simultaneously, shape each other's directive
- User modifies sketch, system suggests new ideas
- Resulting creative product in the end can't be attributed to only one agent
- System doesn't change output when user doesn't modify sketch => can be used as turn-based system
- Menu at top hides configuration, training, settings etc.
- GUI focuses on usability & human-computer interaction, enables co-creativity how user prefers it



References

- Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial intelligence*, 103(1-2), 347-356.
- Boden, M. A. (2007). Creativity in a nutshell. *Think*, 5(15), 83-96.
- Chen, W., & Hays, J. (2018). Sketchygan: Towards diverse and realistic sketch to image synthesis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 9416-9425).
- Chen, W. (2020, July 14). SketchyGAN. Retrieved August 06, 2020, from <https://git.droidware.info/wchen342/SketchyGAN>
- Gao, C., Liu, Q., Xu, Q., Liu, J., Wang, L., & Zou, C. (2020). Image Generation from Freehand Scene Sketches. *arXiv preprint arXiv:2003.02683*.
- Guo, L., Liu, J., Wang, Y., Luo, Z., Wen, W., & Lu, H. (2017, October). Sketch-based image retrieval using generative adversarial networks. In *Proceedings of the 25th ACM international conference on Multimedia* (pp. 1267-1268).
- Jennings, K. E. (2010). Developing creativity: Artificial barriers in artificial intelligence. *Minds and Machines*, 20(4), 489-501.
- Liu, R., Yu, Q., & Yu, S. (2019). An unpaired sketch-to-photo translation model. *arXiv preprint arXiv:1909.08313*.
- Lu, Y., Wu, S., Tai, Y. W., & Tang, C. K. (2018). Image generation from sketch constraint using contextual GAN. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 205-220).
- Osahor, U., Kazemi, H., Dabouei, A., & Nasrabadi, N. (2020). Quality Guided Sketch-to-Photo Image Synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 820-821).
- Sangkloy, P., Burnell, N., Ham, C., & Hays, J. (2016). The sketchy database: learning to retrieve badly drawn bunnies. *ACM Transactions on Graphics (TOG)*, 35(4), 1-12.
- Shen, Y., Gu, J., Tang, X., & Zhou, B. (2020). Interpreting the latent space of gans for semantic face editing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9243-9252).
- Ventura, D. (2017). How to build a CC system paper type: System description paper. ICC.