ISSN 2007-9737

Deep Generative Visual Therapy: GAN-Driven Image Generation for Cognitive Support in ASD

Viridiana Salinas García[1,2,3],Jesús Jaime Moreno Escobar[1,4,5,*], Hugo Quintana Espinosa[1], Mauro Daniel Castillo Pérez[1], Brenda Lorena Flores Hidalgo[1,6], Daphne Sofía González Cano[7]

[1] Instituto Politécnico Nacional,
Escuela Superior de Ingeniería Mecánica y Eléctrica,
Mexico

[2] Instituto Politécnico Nacional,
Escuela Superior de Física y Matemáticas,
Mexico

[3] Instituto Politécnico Nacional,
Centro de Innovación y Desarrollo Tecnológico en Cómputo,
Mexico

[4] Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Mexico

[5] Instituto Politécnico Nacional,
Escuela Superior de Cómputo,
Mexico

[6] Universidad Autónoma Metropolitana,
Departamento de Sistemas,
División de Ciencias Básicas e Ingeniería,
Mexico

[7] Instituto Politécnico Nacional,
Unidad Profesional Interdisciplinaria en Ingeniería,
Mexico

jmorenoe@ipn.mx

Abstract. Children with Autism Spectrum Disorder (ASD) benefit from personalized visual and cognitive stimulation. We present Deep Generative Visual Therapy (DGVT), an interactive system using Generative Adversarial Networks (GANs) to create tailored visual content for stimulating children with ASD. Our method features a custom GAN architecture trained with symbolic and concrete images to generate suitable stimuli for therapy targets such as attention enhancement, visual sequencing, and pattern matching. The system, built on TensorFlow and Keras, is accessible via Google Colab for real-time control and customization by therapists and educators. A series of cognitive games using generated images supports attention, visual discrimination, and memory. Initial assessments with therapists and pilot users showed positive engagement, indicating GAN-generated stimuli can complement traditional cognitive therapy for ASD. This effort connects

generative deep learning with neurodevelopmental treatment to apply adversarial image synthesis in practical sensory and human-centered applications.

Keywords. Generative, adversarial, networks (GAN), autism spectrum disorder (ASD), cognitive stimulation, visual therapy, human-centered AI, deep learning in health.

1 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by symptoms of social communication, repetitive behaviors, and restricted interests [1]. One of the great difficulties in educational and therapeutic efforts is providing interventions that are customizable to the wide range of needs among individuals with ASD.

Visual and cognitive stimulation has become increasingly clear to be an important aspect of the learning experience of children with autism [12]. Recent progress in generative AI, especially in Generative Adversarial Networks (GAN), provides new opportunities to generate personalized visual stimuli that can improve cognitive function and attention [10].

The proposed research responds to the pressing need for sophisticated visual and cognitive stimulation tools for children with Autism Spectrum Disorder.

We present a new algorithm for generating customized dynamic numerical images over static content using the Generative Adversarial Network, which outperforms traditional methods. The main project is called Visual-ASD, and it is an interactive system specifically designed to tailor a variety of visual stimuli to facilitate cognitive testing / rehabilitation in this population. To perform this study, we first conduct a review of the literature on visual stimulation and artificial intelligence methods to identify existing gaps and novel directions.

For that purpose, we design a custom GAN model inference capable of creating high–quality numeral images tailored to the visual perception and recognition abilities of autistic children. The experiments include interactive activities such as memory games and number recall challenges to reinforce lesson themes and develop young viewers' cognitive skills, Figure 1.

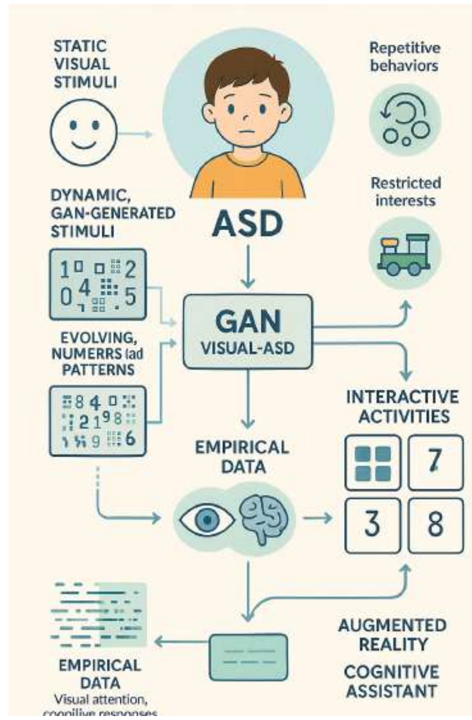

Fig. 1. Overview of the Visual-ASD framework, leveraging GANs to create dynamic, personalized visual stimuli for enhancing cognitive and visual engagement in children with ASD

A well-designed testing scheme is employed to thoroughly assess the effectiveness of the system in objective metrics such as visual attention, cognitive response and accuracy as well as participation.

Further, we take an iterative approach to iteratively refining the GAN model according to empirical feedback that it achieves its success.

Admittedly, this work does not only provide a new therapy tool, but also presents theoretical and implementation understanding on how deep adaptive generative network based representation learning becomes a bridge between deep end-to-end learning and human-driven therapy in the future (towards potential applications to augmented reality and cognitive assistance).

2 Foundations and Prior Work in Generative Visual Stimulation for ASD

The integration of generative artificial intelligence, particularly Generative Adversarial Networks (GANs), has emerged as a promising frontier for the development of adaptive therapeutic tools. This section reviews the technological foundations of GANs, their application in addressing the limitations of conventional methods, and their specific potential to meet the unique needs of individuals with Autism Spectrum Disorder (ASD), thus defining the contribution of our proposed system Visual-ASD.

Introduced by Goodfellow et al. in 2014 [10], the canonical GAN architecture established a powerful adversarial framework for data generation.

As shown in Figure 2, it involves two neural networks—the generator, which generates artificial data from random noise, and the discriminator, which assesses whether the generated samples are simulated real enough as compared to a genuine dataset. This competition encourages the generator to generate more and more realistic results. The inherent capability of GANs to generate rich and realistic images makes it a valuable tool for generating dynamic stimuli that are tailored to the particular perceptual and cognitive profiles of children with ASD [20]. Also, the research design of Lugo-Torres et al. (2023) [17] find that, in GAN-based analyses when data are scarce or not correctly paired (as is common for generative systems for therapeutic purposes), significant improvements of the visual coherence and the semantic content preservation can be achieved through suitable data augmentation strategies. The authors show that processes such as flipping, rotating or inverting pairs of samples enables to stabilize training and diminish common issues like mode collapse and overfitting, which are particularly important if data availability is limited.

Further architectural developments have increased this capability dramatically. Recent work like StyleGAN [13, 14] can synthesise high resolution images with complex, fine-level details. This is not just baby steps; it's a prerequisite for producing educational materials that can be sufficiently customizable to address the range of sensory experiences which exist in autism. However, the use of such powerful generative models in ASD interventions must be taken with care and grounded in evidence that respects both the atypical visual perception as well as cognitive processing characteristics apparent within this population [9]. In contrast, Cerino et al. (2023) [5] propose an investigation that establishes a relationship between symbols extracted from the ARASAAC database and Toki Pona, a minimalist language, in order to evaluate their applicability as pictograms for AAC systems. Using 194 surveys the authors identified which pictograms are consistently interpreted as such, providing a selected set of visual concepts that can be adapted for symbol based communication. This research suggests that picture categorization test-01 visual clarity and iconicity of pictograms are crucial for users with speech or language disabilities, including individuals with ASDintellectual impairments, or motor dysfunction. The results also demonstrate that simplifying the vocabulary, using sound symbols and creating easily interpretable correspondences contribute to enhance the ACC process as well as enable EEG-based devices and solutions for expert users. Similarly, Cortez Martínez et al. (2023) [6], propose a modification of the original Morris Water Maze to an increasing popular mobile VR experience. Pilot test's results seem to demonstrate that participants are able to memorize and learn (at least in the short term) the visual environment, which can help them perform even when distractors are present, thus ensuring the potential value of VR as an available tool for a cognitive assessment. These results collectively contribute to a core assumption of this work, that specifically targeted visual stimulation and therapeutic mediation through the use of images offers much-needed potential as an effective means for communication support and cognitive strategies within neurodiverse populations.

The current approach to traditional cognitive intervention for ASD has relied heavily on static images or rigidly repetitive sequences, with the goal of developing visual attention and motor perception [18]. While these approaches have been found effective, there is growing evidence that

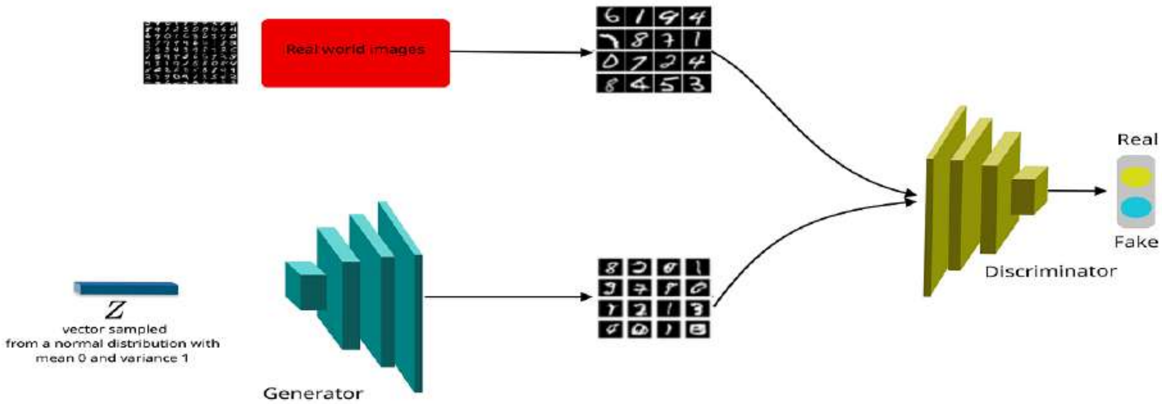

Fig. 2. Generative Adversarial Network Architecture [10]

they may lack the plasticity and dynamism required to optimally accommodate the heterogeneous ASD population. This type of rigidity does not support the many existing individual differences, which may limit therapeutic efficacy and development.

In the field of artificial intelligence applied to mental health, most recent studies have focused on disorders such as depression, anxiety, occupational stress, or automatic identification of personality traits. For example, Bátiz-Beltrán et al. (2024) [4] developed a Spanish linguistic resource for inferring personality traits from text. Meanwhile, Sahoo et al. (2024) [21] conducted research on monitoring critically ill patients through smart medical devices and IoMT infrastructures supported by deep learning models, while Automatic Depression Detection in Social Networks Using Multiple User Characterizations proposes a multimodal fusion model based on GMU to identify signs of depression by combining emotions, demographic traits, and users' thematic content, outperforming traditional text-analysis methods (Valencia-Segura et al., 2023) [23].

Although these studies show important advances in the automatic modeling of psychological indicators, Autism Spectrum Disorder (ASD) remains underrepresented as a primary area of study within AI applied to mental health, highlighting a disadvantage in the development of generative and multimodal model–based tools such as GANs for generating personalized visual stimuli specifically aimed at cognitive and therapeutic intervention in ASD.

Generative models such as GANs enable a new paradigm: instead of selecting existing static materials specific to age, generated content can be automatically adapted to the child in question. By dynamically adjusting the visual curriculum to the child's developmental level and changing preferences, GAN-generated stimuli could overcome the limitations of static content [7, 8]. This adaptive capability is essential for creating a more personalized and engaging therapeutic experience necessary to capture attention and support learning.

The usefulness of GANs goes far beyond simply generating novel images. From a cognitive standpoint, they can be used to construct rich and visually engaging environments explicitly designed to target and improve specific cognitive functions such as working memory and visual discrimination [3, 15]. For example, [22] demonstrated the use of GANs not only for visual stimulation but also to promote cognitive activity through the production of controlled sequences of numbers and abstract patterns, which proved effective in improving sustained attention.

A critical advantage of this generative approach is the capacity for real-time adaptive feedback. Unlike static systems, a GAN-based platform can modulate its output based on user interaction, thus increasing engagement and facilitating stronger cognitive connections with the content. Additionally, the use of personalized visual stimuli has been associated with reductions in anxiety and stress levels in individuals with ASD [2], effectively removing a significant barrier to skill acquisition and enabling a more effective learning environment.

The exploration of neural networks for ASD support is not limited to GANs. Other architectures have significantly contributed to the field. For example, Recurrent Neural Networks (RNNs) have been used by [16] to track the temporal dynamics of cognitive responses to visual stimuli, reporting that dynamic image sequences improved information processing. Similarly, hybrid neural network systems have been proposed for the automatic early detection of ASD-associated traits with high accuracy [19].

Studies such as [11] have explored the integration of deep reinforcement learning to design adaptive models that control stimulus complexity according to the child's real-time performance. This approach not only personalizes the therapeutic process but also helps manage cognitive load, ensuring that challenges always remain within the child's zone of proximal development.

In conclusion, the body of work reviewed here highlights a significant convergence between technological innovation and therapeutic practice. The progression from the foundational GAN framework to its modern, high-fidelity variants and its contrast with conventional static methods outlines a clear trajectory toward highly personalized and adaptive intervention tools. Studies on cognitive engagement and the broader ecosystem of deep-learning applications further validate this direction.

Our proposed system Visual-ASD is built directly on this foundation. It seeks to integrate the high-fidelity generative capabilities of advanced GANs with a structured framework of interactive cognitive exercises and a robust real-time adaptation protocol. In doing so, it aims to address the identified gaps and offer a novel and effective instrument for visual and cognitive stimulation tailored to children with ASD, contributing both to applied therapy and to the theoretical understanding of generative AI in human-centered applications.

A clear practical advantage of GANs is the ability to work efficiently with small datasets. By combining methods such as data augmentation, these networks can produce a wide variety of visual samples from a small corpus of original images.

This is especially important in the ASD field, where obtaining large and unbiased datasets can be difficult. It also allows researchers and others to create personalized stimuli representing certain emotions or social situations, giving children the opportunity to gradually become familiar with what might otherwise be overwhelming. Moreover, GANs hold great potential for diagnostic and assessment methods.

By producing complex visual data that mimic a variety of behaviors and states, they can also be employed to develop more sophisticated methods for investigating individual responsibility. This enables a much deeper and data-driven understanding of a child's individual needs, making it possible for clinicians and therapists to design interventions with an extremely high level of specificity. Finally, the integration of GANs with novel technologies such as augmented and virtual reality paves the way for new therapeutic approaches.

This synergy allows users to be visually immersed in virtual worlds, where children can safely explore and practice social skills and learning activities without the limitations of the real world. These game-like systems can significantly reduce anxiety, improve learning, and promote the generalization of skills—one of the key limitations of existing ASD intervention approaches.

Thus, beyond their virtualization capabilities, GANs provide a multifaceted toolkit for personalization, assessment, and immersion. It is not just about generating content; rather, they constitute an underlying technology that helps create more flexible and responsive systems to support children with ASD.

3 Proposed Methodology and Development Pipeline

3.1 Integrated Technical Stack for Adaptive Therapeutic Stimuli

The applicative solution Visual-ASD is a technical stack designed for the cloud that provides visual adaptive cognitive stimuli with absorbing content that was powered for Generative Adversarial Networks.

Built on it can stand alone as package and once again it grows by being fed into more information. presents an opportunity for the reuse of pretrained GANs, include your own images to run in the app for a mobile yet advanced autism spectrum type intervention would not be missed if something existed that could assist with our now and near future computing facility arrangement/use.

These form the basis for interactive sequencing and recognition exercises which target specific stages of cognitive development.

Google Colab serves as a computation layout that can be scaled to the cloud. Libraries such as `gdown` and `shutil` to get the task of storing the model thus manage effectively (Retrieve by accessing resources) the data and Libraries like *NUMPY* or Utilities for array operations, OS system management. For user input, the system uses *IPython* and `ipywidgets` for interactive exercise interfaces. We visualize the scene with *Matplotlib* for visualization, and we edit images with *PIL* to indicate reactions to the patient with ASD.

This constitutes a compromise between being highly computationally efficient and user friendly, in that it provides for on-line adaptation of the stimuli to individual user properties including responses. The tools have been designed with the aim of being compatible with state-of-the-art generative models and easily deployable in clinical or educational settings, thus fostering the development of personalized therapeutic interactions for children affected by ASD.

3.2 Core Technical Components and Interactive Stimuli Design

The Visual-ASD platform includes two main parts: i) a GAN Training Pipeline, and ii) the Visual-Cognitive Stimulation Game Suite, developed specifically for children with ASD, creating visually realistic images using GANs, which we have allowed to generate highly commended by physicians considering its effectiveness in visual therapy exercises. The GAN Training Pipeline is implemented and the necessary libraries are first imported, and then the generator and discriminator models are defined step by step. Finally, the network is adversarially trained end–to–end, and the generator learns to generate realistic synthetic images, whereas the discriminator adapts its decision boundary based on the generated and true data. The training process ends with model validation and novel imagery generation, making it safe for use in therapeutic applications.

The second module, Visual–Cognitive Stimulation Game Suite, is organized into an initial setup phase and a set of interaction exercises. The configuration includes downloading and setting up pre–trained GAN models, testing that they are functioning correctly, and then importing game libraries. This pretraining step allows the systems to be set up correctly and adapted for each user, so that they work stably and reliably before undergoing further stimulation.

This solution contains five unique interactive games that were created to stimulate different brain functions. These range from Visual Matching, Mouse-Click Based Matching, Interactive Pair Matching to Numerical Sequencing, mathematical exercises, and more. All games utilize the generative power provided by a pre–trained GAN to generate dynamic, adaptive visual content to provide an engaging and personalized therapeutic experience that supports cognitive and visual processing skill development in children with ASD.

3.3 GAN Training Pipeline

The generator structure is intended to generate grayscale images from a noise vector and some conditional label input, Figure 3. The process starts with the concatenation of these inputs, then

ISSN 2007-9737

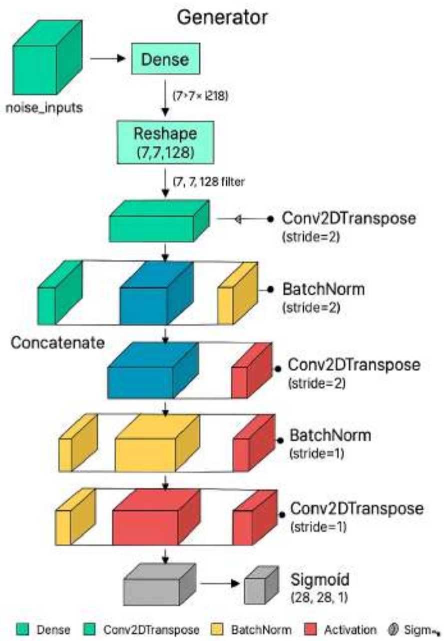

Fig. 3. Architecture of the GAN generator producing numerical images for autism therapy

a dense layer that appends this to the features, and finally reshaping the combined flattened vector into a 3D tensor. This tensor is then up-sampled with several transposed convolutional blocks consists of a *Conv2DTranspose* layer followed by batch normalization and ReLU activation. The number of filters progressively reduces (128 $\rightarrow$ 64 $\rightarrow$ 32) in these blocks, introducing and preserving the feature maps and increasing the spatial resolution. The last block consists of a transposed convolution with a single filter and sigmoid activation function, resulting in a size output $28 \times 28 \times 1$ that represents the generated image. The hierarchical architecture supports the learning of features incrementally and allows the growth of the spatial size of the outputs.

Then, the discriminator scheme is created, which is a critic part of the GAN, Figure 4. It discriminates between *real* vs. *generated* images based on the authenticity. The model has two inputs: i) the image and ii) the corresponding labels that are concatenated after the labels are reshape to the image size. This concatenated input is treated through several layers of convolutions with `LeakyReLU` activations, allowing the extraction of features hierarchically at different levels of abstraction. The output is raster flattened, passed through a fully connected layer with a sigmoid activation function that predicts that the input image is real or not, i.e. the probability between 0 and 1. As one of the important parts of a GAN, the discriminator encourages the generator to generate more realistic images using adversarial training.

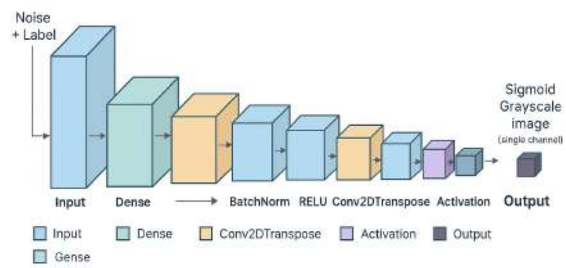

Fig. 4. Architecture of the GAN discriminator producing numerical images for autism therapy

The Model Construction for Training is a method designed to construct and configure the core components of a Generative Adversarial Network, Figure 5. First, the input size and learning parameters are defined and the base discriminator model is built to evaluate the authenticity of the generated images. This discriminator is then configured and compiling using a binary cross-entropy loss function and an *RMSprop* optimizer. The generator model is then constructed to produce images from noise input and labels. A frozen discriminator is also implemented that maintains fixed weights during the adversarial training process. Finally, the adversarial model, combining the generator and the frozen discriminator, is built and compiled to train the generator by assessing the quality of both real and generated images. For this adversarial model, an optimizer with a reduced learning rate is used to ensure stable adversarial training.

Figure 6 shows the Generative Adversarial Network training scheme, Visual-ASD, is specialized in the generation and adaptation of a generator and discriminator for visual and cognitive stimulation. The images from the MNIST repository are resized and form obliquely for training. The discriminator is trained to differentiate real from generated images, when the fake image produced by the generator and the real ones are merged. At the same time, the adversarial network is trained to confuse the discriminator. This procedure trains the generator to obtain even more realistic images. Multiple training steps are used, including periodic assessment and fine-tuning of discriminator and adversarial network performance. The samples produced by the trained models are automatically saved.

The GAN training process consists of creating and evaluating the three main components of the network to construct the generator and discriminator schemes, that is, initializing the generator (G), the discriminator (D) and the adversarial model (A). Once the design architecture has been created. In Python, the `summary()` method is used to validate whether our design is indeed accurate before starting training. The adversarial training loop, controlled by the training function, incrementally trains both parts: the discriminator is trained to discriminate real from generated images in mixed batches, and the adversarial model trains the generator to generate more realistic output to try to fool it. At the end of the training, we estimate the generator's performance by processing noise vectors and numerical labels (0-9) to generate 16 labeled outputs for visualization purposes and to check if a semantically sound digit representation can be generated, proving the model capability in label–to–image mapping.

Finally, Visual-ASD initiates numerical image generation by processing random input noise alongside a hot encoded label (for example, representing the digit 2 or 5), Figure 7. The generator ingests these inputs and synthesizes a corresponding numerical image, which is subsequently rendered in the interface. The inclusion of random noise is critical for introducing diversity in the generated outputs; it acts as a stochastic seed vector, enabling variations in stylistic features (e.g. rendering the digit 5 with distinct curvatures or stroke patterns). Without

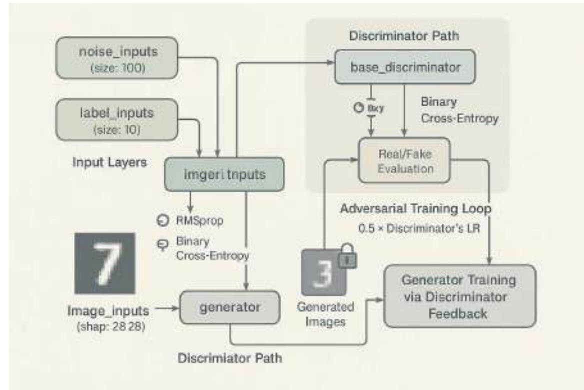

Fig. 5. Architecture of the GAN training pipeline showing input processing (noise vector and one-hot labels), generator and discriminator networks, and adversarial training loop with gradient flow

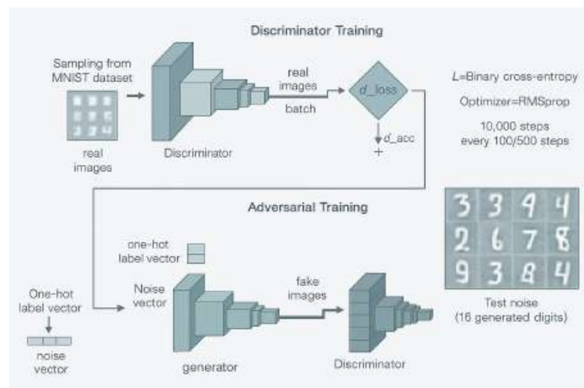

Fig. 6. Adversarial training dynamics of the GAN: Discriminator learns to distinguish real (MNIST) and generated images, while the generator (via adversarial loop) improves by fooling the frozen discriminator

this noise, the generator would deterministically produce identical outputs for each label, severely limiting its capacity for expressive and novel image synthesis.

3.4 Visual–Cognitive Stimulation Game Suite

The Visual-ASD project uses a generative model to create artificial numbers and embed them in an interactive environment for visual and cognitive stimulation of children with autism, Figure 8. These activities range from a

ISSN 2007-9737

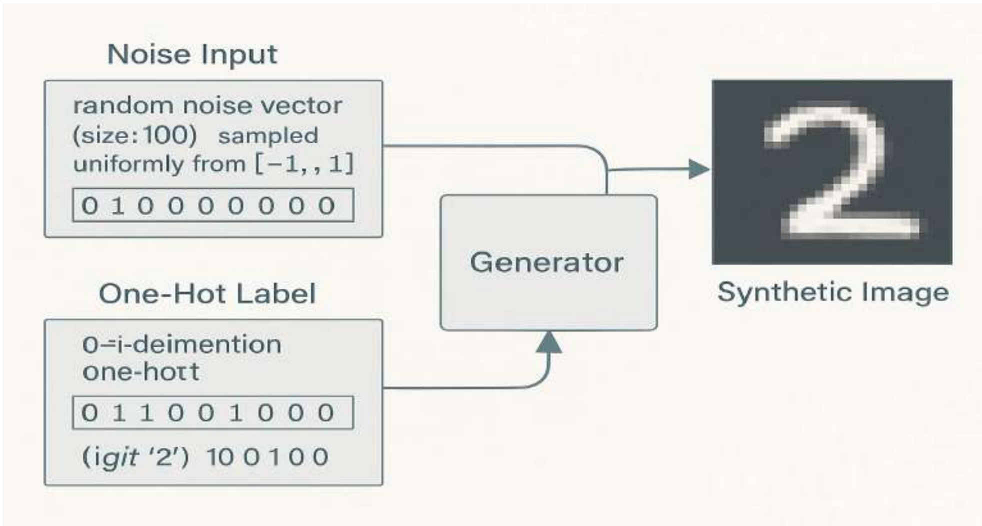

Fig. 7. Fake digit generation via trained GAN: Noise vector and one-hot label are processed by the generator to synthesize a digit image

matching game, in which children match numbers presented in array form (e.g., *How many?*). to numerical sequencing and arithmetic tasks that involve sequence repetition and computation with graphically displayed digits, respectively. The generated images are incorporated into interactive *widgets*, allowing children to interact and consume content while learning through a visual experience.

The Visual-ASD protocol includes a grid with synthetically obtained numeric images, and children play the *visual matching game*, in which they are required to find identical numbers within a matrix of digits.

This task builds pattern recognition and attention to detail using moving visual stimuli, tailored to the sensory processing preferences of autistic learners. As players learn and improve, the system can adjust complexity that supports progressive skill acquisition in a managed but elastic manner.

Memory-click game of the generator numbers, a mouse-driven level dedicated to improve interactivity, as players have to reveal pairs in a generated numbers by clicking on cells. The interface logs the skill and response time, instantaneously rewarding correct responses to aid in learning.

This method not only promotes working memory, but also hand–eye coordination, and includes some interactive components to keep the patient engaged during their treatment.

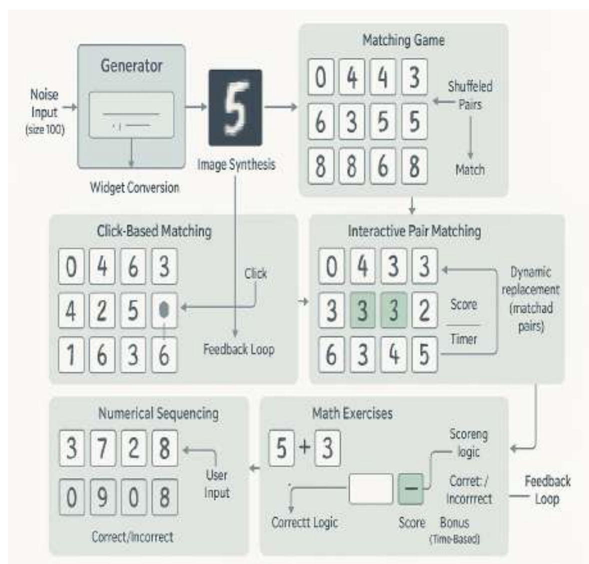

Fig. 8. Architecture of the GAN Visual-Cognitive Stimulation Game Suite: Synthetic digit generation drives five interactive games (matching, sequencing, arithmetic) with real-time feedback and adaptive difficulty for autism therapy

The *interactive pair matching exercise* adds an element of difficulty by adding a time challenge and a visual reinforcement packet to find pairs. The kids discover pairs of hidden numbers on a grid; the game monitors progress and adjusts the level of difficulty according to skill. It helps develop executive function skills (such as planning and impulse control) and minimize frustration by automatically adjusting the difficulty level.

For *sequential memory training*, the model uses a number sequence game in which children watch and reproduce increasingly long sequences of digits read from the model. Scoring Feedback involves point tallies for accuracy and speed and subtracts the points for errors, thereby availing a balanced feedback system in driving improvement. The purpose of this drill is to develop short-term memory and auditory-visual integration skills.

Finally, the *arithmetic challenge module* uses generated numbers to create interactive math problems ranging from simple addition to subtraction. Attendants solve equations presented as visual numbers, while given instant performance metrics that highlight their strengths and weaknesses. The system is designed to engage the mind in that you have to solve problems, answer math questions correctly, and in turn close relationships between mathematical dialects produced by children, both domain specific and global.

4 Experimental Evaluation

4.1 Functional Testing of GAN-Based Numerical Image Synthesis

The experimental testing consists of two main processes: testing with the pre–trained model and targeted fake image generation. The former demonstrates the generator's ability to generate all digits (0-9) using noise vectors and one-hot labels, and the latter address a detail of generating specific numbers that is necessary for physical therapy practice. The two mechanisms use random noise to create diversity in the visual output, such that same-label inputs generate diverse styled image representations. This variability is necessary to keep children with autism challenged and engaged during interactive activities, Figure 9.

We start with the validation of the pre–trained model to ensure that GAN is functioning. The approach is quite simple and just generates random noise inputs attached with one-hot encoded labels to create 16 fake images for every digit (0-9) as input. Figure 10 outputs are displayed to understand whether the generator could generate reasonable numerical representations on all categories. For the targeted fake generation, specific digits (for example, '2' or '5') are generated by adding noise to the corresponding hot labels, generating personalized visual stimuli for therapeutic exercises. The key to this approach is then the usage of random noise vectors, allowing one to impose noise on stylistic features (i.e., curvature and stroke thickness) for identical labels. Without this randomness, the generator would generate a static output, which would limit its usefulness in generating interesting cognitive stimuli. This variability is essential for therapies targeting autism, with dynamic visual contexts being beneficial to adaptive learning and prolonged engagement.

4.2 Assessment of Therapeutic Gaming Interventions

VISUAL-ASD uses an image creator to produce a matching game that operates on visual stimuli. The activity generates visual numbers or symbols that the children must match correctly and aims to increase both visual and cognitive stimulation in children with autism. Using cutting-edge technology, the game grows to fit user needs, creating a personalized therapeutic experience. As we show in Figure 11, the game produces ten pairs of numerical stimuli for visual matching, using appealing, GAN-produced stimuli to interest the players in pattern recognition and detail focus.

The second game of Visual-ASD suite employs a generative network to produce numerical images shown on screen as a grid with clickable buttons. Two pairs of identical numbers are revealed by clicking the buttons that children with ASD must match (see Figure 12). In-game learning is constantly assessed and feedback offered

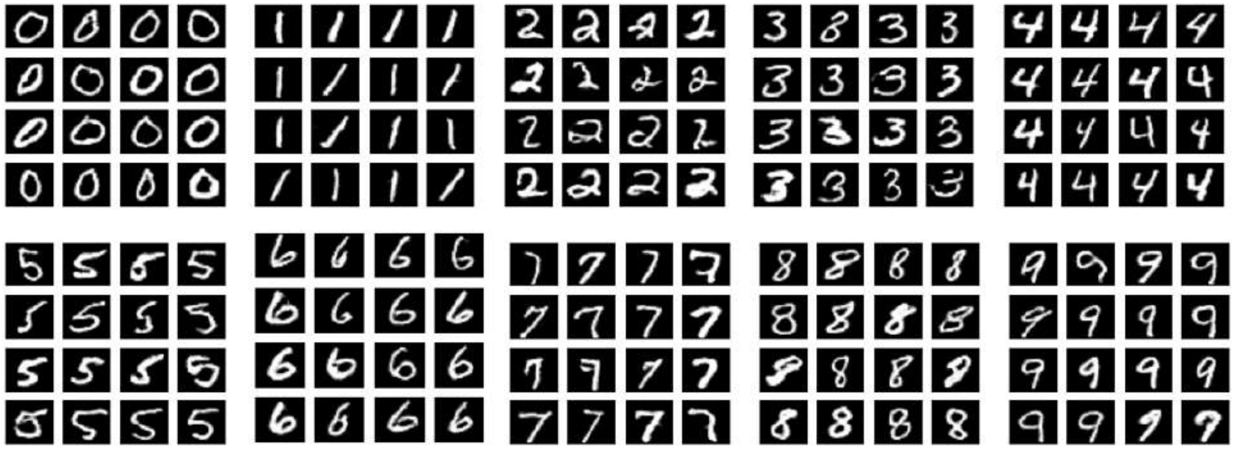

Fig. 9. Pre-trained Model Validation: Synthetic digit generation (0-9) via GAN using noise vectors and one-hot encoded labels

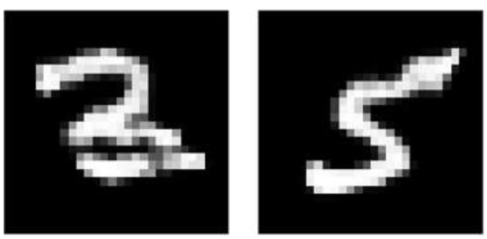

Fig. 10. Targeted Digit Synthesis: Stochastic generation of digits 2 and 5 demonstrating output diversity from identical labels

instantly, to reinforce your understanding. This is a memory skills and brain training game that will challenge your brain and stimulate its sensory activities, improve logic, mathematical equations, and cognitive functions along with spatial intelligence and memory tests. The game ends when a player makes all ten matches, the cumulative time taken is shown on as a performance parameter.

Visual-ASD also has a matching game that uses a generative image network to generate and obfuscate pairs of digits. This game involves a grid of individual numeric images that are initially hidden beneath clickable buttons. The players push the buttons and try to find the matching numbers. After a successful match, the pair is

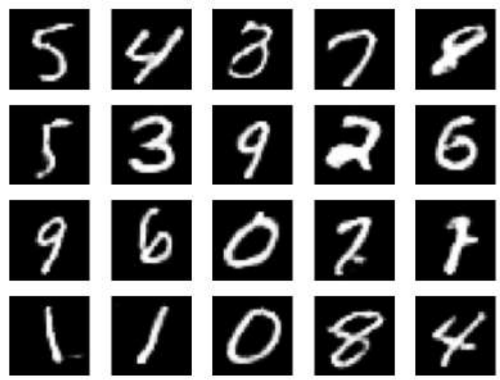

Fig. 11. Game 1: Visual Matching Game. Grid of synthetically generated numerical pairs for associative learning

removed from the board, and once all 10 pairs have been matched, the game ends (the grid will be empty). The system logs every interaction (successes, failures, and overall task time), which is itself a form of feedback in real time that helps to stimulate visual and cognitive behavior when performing an activity for the child with autism.

The sequence game is a GAN-based memory and cognitive processing promoting game. Players are shown a sequence of numerals produced by the model, which they memorize and then replay in turn. As the player progresses, the sequence

Fig. 12. Game 2: Mouse-Click Based Visual Matching. Interactive grid for memory and matching task performance tracking

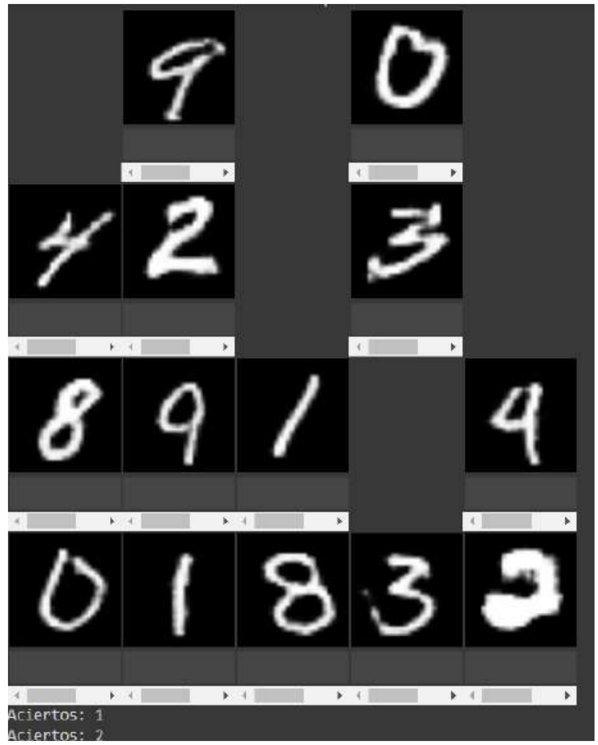

Fig. 13. Game 3: Interactive Pair Matching. Dynamic grid with hidden numerical pairs for memory and matching exercises

length gradually augments, premeditating memory retention and attentional load. The platform assesses performance in terms of accuracy and speed, offering instantaneous feedback for cognitive enhancement and intelligent learning.

Visual-ASD software also includes a Arithmetic Exercises, developed for providing interactive practice of written arithmetic.

This module performs basic math operations, in the form of addition and subtraction, on numerical images produced from a generative network.

Users try to solve the shown matrix problems and system automatically computes their performance score according to correctness of answers.

The application includes real-time tracking of performance status and is designed to increase active and cognitive participation in children with ASD.

5 Conclusions

In this work, we introduced Visual-ASD, a new generation of interactive gaming combined with the framework of Generative Adversarial Networks that directly targets children suffering from Autism Spectrum Disorder (ASD) in terms of both visual and cognitive stimuli. The project showcases the use of cutting-edge deep learning techniques, such as GAN, to build dynamic, adaptive, and personalized therapeutic instruments. Using a variety of interesting graphical material, it transcends from an ordinary static intervention technique to implement the scale problematic and provide open-source access to neurodevelopmental treatment.

Technically, the combination of *TensorFlow, Keras*, and an interactive programming library such as `ipywidgets` in the Google Colab cloud-based app was helpful to create a powerful yet

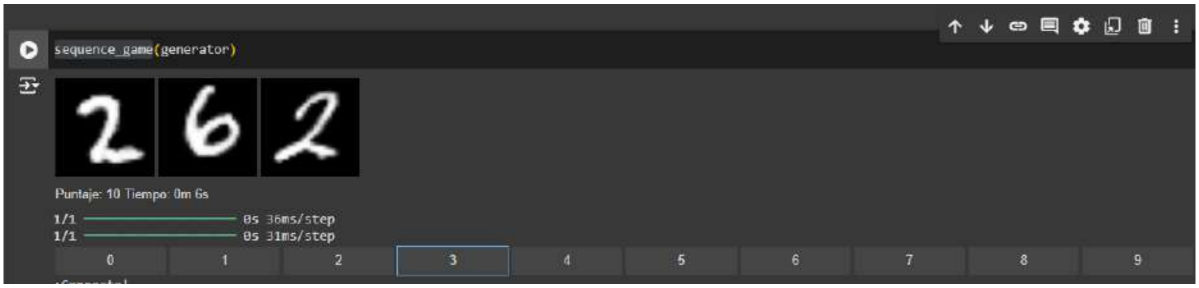

Fig. 14. Game 4: Numerical Sequencing. Adaptive memory task with progressively increasing sequence length and real-time scoring

Fig. 15. Game 5: Arithmetic Exercises. Interactive math challenges using GAN-generated numerals with real-time scoring feedback

user-friendly interface. The GAN training pipeline was able to generate high-quality numerical images, and the game suite converted these into exercises aimed at memory, attention, and executive function. Using noise vectors provided diversity in the output, which was essential to keep the patient involved and prevent boredom when clinicians used the application over long periods.

From a social perspective, Visual-ASD is a major step toward the development of AI-driven inclusive interventions. Each game, related to visual matching, memory monitoring, sequencing, and arithmetic, was intended to address particular cognitive abilities in a manner that respects the sensory needs of autistic students. Its real-time adaptability and interactive feedback is consistent with what is well known about best practice in autism therapy, that it is all about unique learning pathways and helping educators/therapists help themselves.

Future efforts will conduct clinical validation with longitudinal studies, as well as extend to more general types of stimuli (e.g., objects and emotions), including integration within augmented reality for immersive educational scenarios. The framework is open source to foster collaboration and modification in therapeutic or educational domains. Lastly, Visual-ASD demonstrates that generative AI has a role in developing technologies with social impact at the convergence of machine learning and human-centered care.

Acknowledgment

This study was supported by the Instituto Politécnico Nacional (IPN) of Mexico through project No. 20250776 under the project titled *Neurodecodificación de Preferencias Alimentarias: Estrategias de Prevención Basadas en Inteligencia Artificial para la prevención de la Epidemia de Obesidad-Diabetes en México*, funded by the Secretaría de Investigación y Posgrado, Comisión de Operación y Fomento de Actividades Académicas, and by the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) of Mexico. This study is also part of the 2025 Call for InterInstitutional Collaboration Projects IPN-UAM-UAEMÉX under the project titled *Desarrollo de una Aplicación de Inteligencia Artificial para el seguimiento de contaminantes, salud, y Análisis de Factores Determinantes para el Estado de México*, by through Project No.

IPCC-008-2024. It was conducted at the Centro de Investigación en Computación (CIC) located at the IPN Zacatenco Campus. This article is part of the doctoral thesis of Viridiana Salinas, entitled *Metodología de sistemas suaves en la detección de trastorno del espectro autista a partir de imágenes de resonancia magnética*, supervised by Dr. Jaime Moreno. It is also important to highlight that this work is part of the Curricular Professional Internship of Daphne González from the B.A. in Artificial Intelligence, conducted at the Computational Cognitive Sciences Laboratory of the Center for Computer Research. Furthermore, we note that a portion of this work was developed by Mauro Castillo as a product of his master thesis entitled *Modelo Sistémico para Explicar el Comportamiento de los Consumidores durante la Toma de Decisiones, Basado en Inteligencia Artificial*, supervised also by Dr. Jaime Moreno.

References

1. **Association, A. P. (2013).** Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association. DOI: `10.1176/appi.books.9780890425596`.

2. **Baron-Cohen, S. (2013).** Understanding Other Minds. Oxford University Press, Incorporated, Oxford, 3rd ed. edition. Description based on publisher supplied metadata and other sources.

3. **Bau, D., Zhu, J.-Y., Strobelt, H., Zhou, B., Tenenbaum, J. B., Freeman, W. T., Torralba, A. (2018).** Gan dissection: Visualizing and understanding generative adversarial networks. DOI: `10.48550/ARXIV.1811.10597`.

4. **Bátiz-Beltrán, V. M., Barrón-Estrada, M. L., Zatarain-Cabada, R., Roldán-Arana, J. I. (2024).** Creation of a corpus in spanish for the recognition of personality traits. Computación y Sistemas, Vol. 28, No. 3. DOI: `10.13053/cys-28-3-4619`.

5. **Cerino Jiménez, R., Pinto Avendaño, D. E., Vergara Limon, S. (2023).** Pictographic representation of the toki pona language for use in augmentative and alternative communication systems. Computación y Sistemas, Vol. 27, No. 2. DOI: `10.13053/cys-27-2-4418`.

6. **Cortez Martinez, C. R., Arámburo-Lizárraga, J., Cortez Álvarez, C. R., Ramírez Herrera, M. A., Mendoza-Magaña, M. L. (2023).** Development of a mobile virtual reality application based on morris water maze for the evaluation of spatial memory in human beings. Computación y Sistemas, Vol. 27, No. 4. DOI: `10.13053/cys-27-4-3388`.

7. **Dunmore, A., Jang-Jaccard, J., Sabrina, F., Kwak, J. (2023).** A comprehensive survey of generative adversarial networks (gans) in cybersecurity intrusion detection. IEEE Access, Vol. 11, pp. 76071–76094. DOI: `10.1109/access.2023.3296707`.

8. **Fahimi, F., Dosen, S., Ang, K. K., Mrachacz-Kersting, N., Guan, C. (2021).** Generative adversarial networks-based data augmentation for brain–computer interface. IEEE Transactions on Neural Networks and Learning Systems, Vol. 32, No. 9, pp. 4039–4051. DOI: `10.1109/tnnls.2020.3016666`.

9. **Frith, U. (2008).** Autism: Explaining the Enigma. Blackwell, Malden, Mass. [u.a.], 2. ed., [nachdr.] edition. Includes bibliographical references and index.

10. **Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. (2014).** Generative adversarial networks. DOI: `10.48550/ARXIV.1406.2661`.

11. **He, K., Zhang, X., Ren, S., Sun, J. (2016).** Identity mappings in deep residual networks. DOI: `10.48550/ARXIV.1603.05027`.

12. **Hodges, H., Fealko, C., Soares, N. (2020).** Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation.. Translational pediatrics, Vol. 9, pp. S55–S65. DOI: `10.21037/tp.2019.09.09`.

13. **Karras, T., Laine, S., Aila, T. (2018).** A style-based generator architecture for generative adversarial networks. DOI: `10.48550/ARXIV.1812.04948`.

14. **Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T. (2019).** Analyzing and improving the image quality of stylegan. DOI: 10.48550/ARXIV.1912.04958.

15. **Lembono, T. S., Pignat, E., Jankowski, J., Calinon, S. (2020).** Learning constrained distributions of robot configurations with generative adversarial network. DOI: 10.48550/ARXIV.2011.05717.

16. **Liu, Y. H., Zhu, J., Constantinidis, C., Zhou, X. (2021).** Emergence of prefrontal neuron maturation properties by training recurrent neural networks in cognitive tasks. iScience, Vol. 24, No. 10, pp. 103178. DOI: 10.1016/j.isci.2021.103178.

17. **Lugo-Torres, G., Valdez-Rodríguez, J. E., Peralta-Rodríguez, D. A., Calvo, H. (2023).** Comparison of transfer style using a cyclegan model with data augmentation. Computación y Sistemas, Vol. 27, No. 4. DOI: 10.13053/cys-27-4-4775.

18. **Martelli, M. E., Gigliotti, F., Giovannone, F., Lentini, G., Manti, F., Sogos, C. (2025).** Developmental patterns in autism and other neurodevelopmental disorders in preschool children. Children, Vol. 12, No. 2, pp. 125. DOI: 10.3390/children12020125.

19. **Radhakrishnan, M., Ramamurthy, K., Shanmugam, S., Prasanna, G., S, V., Y, S., Won, D. (2024).** A hybrid model for the classification of autism spectrum disorder using mu rhythm in eeg. Technology and Health Care, Vol. 32, No. 6, pp. 4485–4503. DOI: 10.3233/thc-240644.

20. **Rivard, M., Terroux, A., Parent-Boursier, C., Mercier, C. (2014).** Determinants of stress in parents of children with autism spectrum disorders. Journal of Autism and Developmental Disorders, Vol. 44, No. 7, pp. 1609–1620. DOI: 10.1007/s10803-013-2028-z.

21. **Sahoo, S., Panigrahi, C. R., Pati, B. (2024).** Iomt-enabled smart healthcare models to monitor critical patients using deep learning algorithms: A review. Computación y Sistemas, Vol. 28, No. 4. DOI: 10.13053/cys-28-4-4915.

22. **Taigman, Y., Polyak, A., Wolf, L. (2016).** Unsupervised cross-domain image generation. DOI: 10.48550/ARXIV.1611.02200.

23. **Valencia-Segura, K. M., Escalante, H. J., Villasenor-Pineda, L. (2023).** Automatic depression detection in social networks using multiple user characterizations. Computación y Sistemas, Vol. 27, No. 1. DOI: 10.13053/cys-27-1-4540.

Article received on 24/11/2025; accepted on 04/01/2026
**Corresponding author is Jesús Jaime Moreno Escobar.*