

# GeometricShapes: Exergame basado en procesamiento de imágenes para niños con trastorno del espectro autista (TEA)

Solé, Matias  
Universidad Nacional del Litoral  
Facultad de ingeniería y ciencias  
hídricas  
Santa Fe, Argentina  
matisole123@gmail.com

Maldonado, Nahuel  
Universidad Nacional del Litoral  
Facultad de ingeniería y ciencias  
hídricas  
Santa Fe, Argentina  
nahuemaldonado98@gmail.com

Pratto, Tobias  
Universidad Nacional del Litoral  
Facultad de ingeniería y ciencias  
hídricas  
Santa Fe, Argentina  
tobipratto12@gmail.com

**Resumen**—Este informe presenta el desarrollo y la implementación de un exergame accesible y económico diseñado para estimular las habilidades cognitivas y motoras de niños con trastorno del espectro autista (TEA). El juego, se basa en la detección de formas y colores mediante la utilización de una cámara web. Ofrece una experiencia fluida y en tiempo real, buscando la identificación de paletas de colores, mostradas por los niños y previamente sugeridas en pantalla, en formato audiovisual, por la aplicación. Los métodos que utiliza para conseguirlo son la segmentación de color en el espacio HSV para obtener solo aquellos tonos necesarios para analizar, la detección de bordes con el algoritmo de Canny y, por último, la detección de contornos y aproximación de polígonos para el etiquetado. El juego asiste al usuario con ayudas gráficas en los casos que el niño presente dificultades en la asociación de los colores y formas.

**Palabras clave**—exergame, TEA, segmentación de color, detección de bordes, procesamiento de imágenes.

## I. INTRODUCCIÓN

Los niños con trastorno del espectro autista presentan grandes desafíos en áreas como la atención, la percepción sensorial, la comunicación y habilidades motoras [1]. Estos desafíos pueden dificultar el aprendizaje y la interacción con su entorno, lo que afecta su desarrollo integral y calidad de vida [2]. Una de las áreas más afectadas en los niños con TEA es la capacidad de procesar información visual y sensorial, habilidades fundamentales para el desarrollo cognitivo [3].

La prevalencia del trastorno del espectro autista (TEA) ha generado la necesidad de herramientas terapéuticas innovadoras que ayuden en el desarrollo de habilidades cognitivas y motoras. Los exergames, que combinan ejercicio físico y elementos de juego interactivo, se han destacado como una metodología prometedora para la intervención en niños con TEA [4]. Estos juegos interactivos no solo fomentan la actividad física, sino que también pueden ser diseñados para trabajar en habilidades específicas, como la identificación de formas y colores, la coordinación mano-ojo y la planificación motora. Además, los exergames proporcionan una plataforma atractiva y motivadora que

puede captar la atención de los niños de manera más efectiva que las terapias tradicionales [5].

Existen numerosos estudios enfocados en la utilización de estas tecnologías para ayudar a los niños con autismo a mejorar sus habilidades motoras [6,7,8]. FroggyBobby [6], es una exergame que a través del juego, invita a los niños con autismo a atrapar objetos virtuales con la lengua de una rana que es controlada por los movimientos de los brazos, convirtiéndose en una herramienta que ayuda a mantener la concentración y coordinación visio-motriz. Otras investigaciones han explorado el uso de tecnologías avanzadas, como la realidad aumentada y los dispositivos de seguimiento de movimiento, para ayudar en la intervención terapéutica de niños con TEA [9]. Sin embargo, estas tecnologías a menudo son costosas y no accesibles para todas las familias o instituciones educativas.

Es por esto, que este trabajo propone el desarrollo de un exergame accesible y económico que utiliza una cámara web para la captura de imágenes y algoritmos de procesamiento digital para la detección de formas y colores. El juego invita a los niños a identificar y levantar una paleta con una forma específica de un color determinado, proporcionando retroalimentación visual y auditiva para fomentar el aprendizaje y la interacción. La implementación de este exergame ofrece una solución práctica y efectiva para mejorar las habilidades cognitivas y motoras de los niños con TEA.

## II. PROPUESTA DE SOLUCIÓN

### A. Exergame

En esta primera versión, el juego trabaja con tres colores (rojo, verde y azul) y tres formas (cuadrado, círculo y triángulo), los cuales son seleccionados de manera aleatoria. Aunque es posible agregar más colores y formas, se ha optado por limitar la variedad para adaptar la dificultad del juego a niños que están comenzando a familiarizarse con las formas y colores de los objetos que los rodean.

El flujo seguido por el exergame, desde que inicia hasta que termina el juego (Figura 1).

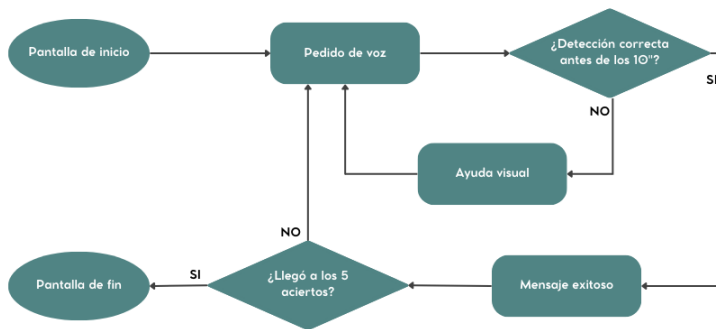


Figura 1. Diagrama de flujo del exergame

Al iniciarse el juego, se presenta una pantalla de inicio (Figura 2a) y el niño, o quien lo acompañe, presiona el botón de comenzar, dando inicio a la interacción. Mediante texto y voz, el exergame invita al niño a levantar una paleta de un color y una forma específica, seleccionados aleatoriamente. En esta fase, se lleva a cabo la detección de la paleta levantada por el niño. Si la paleta no es detectada dentro de un período de 10 segundos, el juego presenta un pictograma como ayuda adicional y vuelve a reproducir la solicitud de voz (Figura 2b). Cuando el niño levanta la paleta correcta, el juego lo felicita con un mensaje para estimular su motivación y participación activa. Este ciclo se repite hasta que el niño alcanza cinco aciertos, momento en el cual el juego finaliza mostrando una pantalla de felicitaciones por completar el juego (Figura 2c).

Para maximizar la concentración del niño, el juego emplea un diseño visual minimalista, evitando distracciones visuales innecesarias. Además, se utiliza una fuente de letra diseñada específicamente para personas disléxicas, lo que facilita la lectura y reconocimiento de las palabras. También se ha priorizado la claridad en la comunicación verbal ya que la voz que guía al niño fue seleccionada cuidadosamente por su claridad y facilidad de comprensión. Asimismo, se han incorporado una serie de pictogramas que actúan como soportes visuales, ayudando al niño a entender mejor las instrucciones. Estos elementos en conjunto aseguran que el juego sea accesible y fácil de seguir para los niños con TEA.

### B. Procesamiento y análisis de imágenes

En esta sección se describen los pasos de procesamiento y análisis de imágenes que realiza el exergame

para cumplir con el objetivo del juego (Figura 3). Este proceso es fundamental para interpretar las acciones del usuario y proporcionar la retroalimentación adecuada, y se compone de 4 fases secuenciales:

- La primera fase es la adquisición de imágenes, donde se captura la imagen en tiempo real utilizando una cámara web. La calidad de esta imagen es importante, ya que se convierte en la base para todo el procesamiento posterior.
- A continuación, la imagen adquirida se somete a una segmentación de color. En esta etapa se aplica un filtro que permite aislar los objetos del color buscado y eliminar la información de color irrelevante.
- La siguiente fase es la detección de bordes donde la imagen segmentada se procesa para resaltar los bordes de los objetos presentes, lo que facilita la posterior identificación de formas.
- Finalmente, la fase de detección de formas localiza y analiza los contornos para identificar la forma buscada en la imagen basándose en las características de los polígonos derivados de cada contorno. La secuencia de procesos se pueden observar en el siguiente esquema.

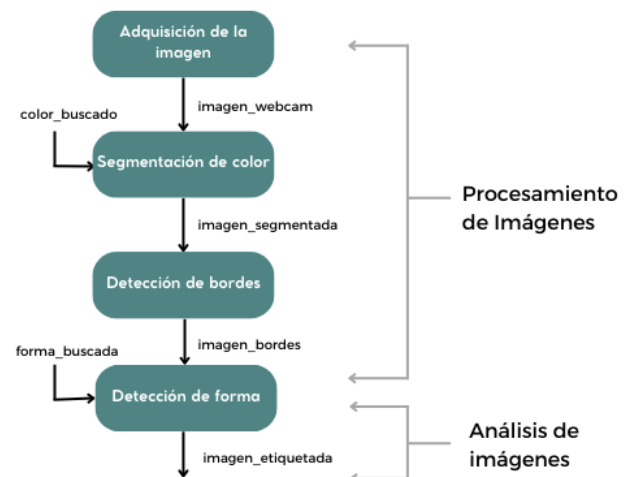


Figura 3. Diagrama de flujo del procesamiento de imágenes

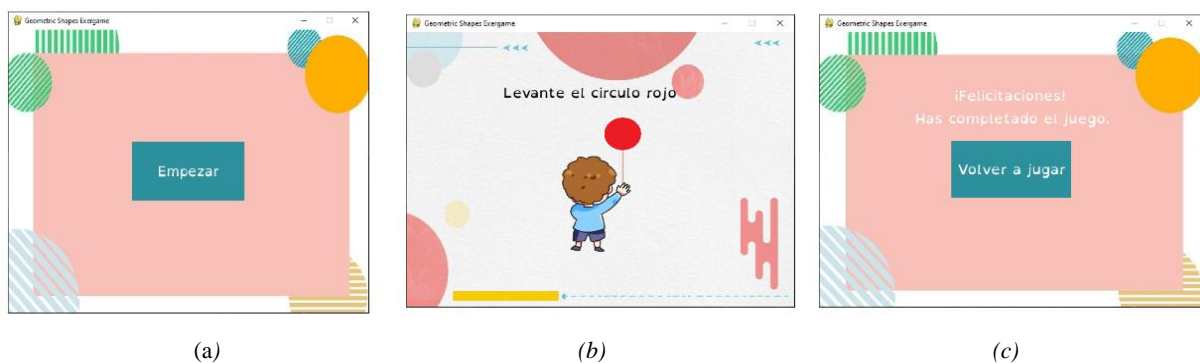


Figura 2. Captura de pantallas: (a) Pantalla de inicio; (b) Pantalla de juego; (c) Pantalla de fin

### 1) Adquisición de la imagen

El proceso de adquisición de la imagen es la primera fase de este flujo, y proporciona la base sobre la cual se realizan todas las operaciones de procesamiento subsecuentes. Este proceso comienza con la captura en tiempo real de imágenes a través de la cámara web integrada en el sistema. El juego configura la cámara para operar a una resolución de 640x480 *pixels*, y a una tasa de captura de 30 *frames* por segundo. Esta configuración asegura que las imágenes capturadas no ralenticen el procesamiento, pero mantengan la calidad suficiente para una detección certera. Como salida de esta etapa, se obtiene cada *frame* de la webcam representado como una imagen en el formato RGB (Figura 4).



Figura 4. Adquisición de la imagen

### 2) Segmentación de color

Por defecto la información del color de cada píxel de una imagen se almacena mediante los canales RGB, correspondientes a la proporción de rojo, verde y azul del píxel en cuestión. Para llevar a cabo la segmentación de color se suele optar por el modelo HSV, el cual cuenta con canales para el tono (*Hue*), saturación (*Saturation*) e intensidad (*Value*). Este modelo separa la información cromática de la de intensidad, lo que permite aislar en la imagen las áreas con un tono similar al deseado.

Una solución para la segmentación de color consiste en medir la distancia euclídea entre los píxeles de la imagen y un punto representativo en el espacio HSV de la paleta, descartando aquellos píxeles cuya distancia supere un umbral predeterminado. No obstante, esta metodología es inadecuada para ambientes con diferentes niveles de iluminación, ya que permitir mayor variación en la intensidad también admitiría tonos alejados del color buscado.

Por lo tanto, resulta apropiado definir rangos personalizados para cada canal del modelo HSV (Tabla 1). De esta forma la segmentación filtra los puntos HSV donde al menos uno de sus canales se encuentra fuera del rango establecido para ese canal. En este caso se propone utilizar rangos de 30 valores para el tono, 150 para la saturación y 200 para la intensidad. Esta configuración permite una segmentación más precisa y consistente bajo diversas condiciones de iluminación, asegurando que sólo los píxeles con los tonos deseados sean seleccionados.

Color	Rango de tono (H)	Rango de saturación (S)	Rango de Intensidad (V)
Rojo	[165-15]	[105-255]	[55-255]
Verde	[35-65]	[105-255]	[55-255]
Azul	[90-120]	[105-255]	[55-255]

Tabla 1. Rangos de valores HSV por colores

En conclusión esta etapa permite pasar de la imagen obtenida de la webcam a una imagen enmascarada donde se mantienen los píxeles con valores HSV similares a los valores representativos de la paleta y el resto se vuelve 0 (Figura 5).

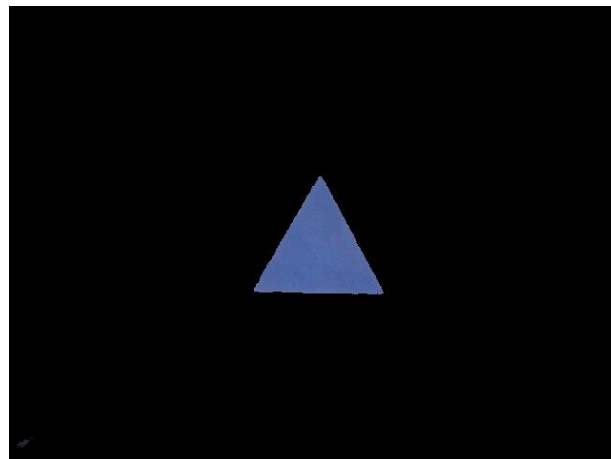


Figura 5. Segmentación de tonalidad azul

### 3) Detección de bordes

Entre los diferentes algoritmos disponibles para la detección de bordes, el algoritmo de Canny se destaca por su robustez y eficiencia. A continuación, se describen los pasos principales de este algoritmo y su aplicación en este contexto.

El primer paso es la reducción de ruido en la imagen, ya que el ruido puede causar la detección de bordes falsos. Para ello, se utiliza un filtro pasabajos, que suaviza la imagen y reduce la variabilidad en los valores de los píxeles. Este filtro aplica una convolución que atenúa las variaciones abruptas.

Una vez que la imagen ha sido suavizada, el siguiente paso es calcular los gradientes de intensidad en la imagen. Estos gradientes indican la dirección y magnitud de los cambios de intensidad en los píxeles. Se utilizan filtros de Sobel para derivar estos gradientes en las direcciones horizontal ( $G_x$ ) y vertical ( $G_y$ ). La magnitud del gradiente  $G$  y su dirección  $\theta$  se calculan utilizando las siguientes fórmulas:

$$G = \sqrt{G_x^2 + G_y^2}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

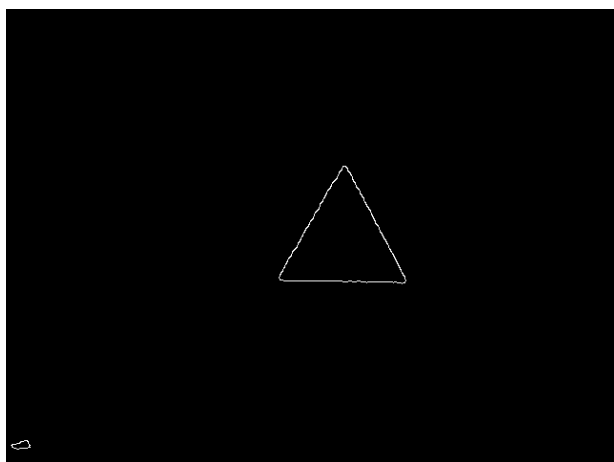
Después de obtener la magnitud y dirección de los gradientes, se aplica la supresión de no-máximos para afinar los bordes. Este paso implica recorrer cada píxel y suprimir aquellos que no son máximos locales en la dirección del gradiente. En otras palabras, si un píxel no tiene el valor más

alto de gradiente en su vecindad a lo largo de la dirección del gradiente, su valor se establece en cero.

La doble umbralización es el siguiente paso, donde se aplican dos umbrales, uno alto y otro bajo. Los píxeles con gradientes superiores al umbral alto se consideran bordes fuertes, mientras que los píxeles con gradientes entre los dos umbrales se consideran bordes débiles. Los píxeles con gradientes por debajo del umbral bajo se descartan.

En el paso final, se realiza un análisis de conectividad para asegurar que los bordes débiles conectados a los bordes fuertes se mantengan, mientras que los bordes débiles no conectados se descartan. Esto ayuda a eliminar falsos positivos y asegura que los bordes detectados representen verdaderas transiciones en la imagen.

En conclusión, la detección de bordes mediante el algoritmo de Canny transforma la imagen segmentada en una representación binaria donde los bordes de las formas de interés están claramente delineados (*Figura 6*).



*Figura 6. Detección de bordes*

#### 4) Detección de formas

Después de detectar los bordes en la imagen segmentada, se busca identificar las formas presentes en la imagen. Este proceso es esencial para que el sistema pueda reconocer y validar las paletas mostradas por el niño. A continuación, se describe el proceso de detección de formas de manera general.

Primero, se identifican los contornos en la imagen utilizando el algoritmo de Suzuki y Abe [10]. Este algoritmo sigue un método conocido como "border following" o "border tracing" para detectar contornos en imágenes binarias. Inicia buscando el primer píxel de borde en la imagen y, una vez encontrado, sigue el contorno registrando las coordenadas de los píxeles de borde en una dirección específica, hasta regresar al píxel inicial. Este algoritmo asegura una detección precisa de los límites de los objetos presentes en la imagen.

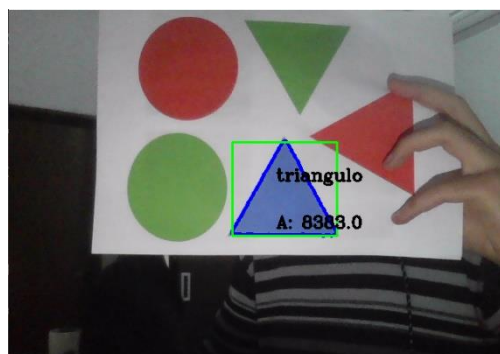
Luego, para evitar el procesamiento de formas irrelevantes o demasiado pequeñas, se filtran los contornos en función de su área. Es decir, que solo se consideran aquellos objetos que superan un tamaño mínimo predefinido. En este caso se decidió fijar este tamaño mínimo en 1000 píxeles, ya que es aproximadamente el área ocupada por una paleta al ubicarse a 2.5 metros de la cámara.

Cada contorno se analiza para aproximar su forma a una figura poligonal con un número específico de vértices. Esta aproximación ayuda a simplificar y clasificar la forma del objeto. La clasificación de cada una de las formas posibles de las paletas depende de los siguientes criterios:

- **Triángulo:** Si la figura aproximada tiene tres vértices, se clasifica como un triángulo.
- **Cuadrado:** Si la figura tiene cuatro vértices, se calcula la relación entre el ancho y el alto de la figura. Si esta relación está cerca de uno, se clasifica como un cuadrado; de lo contrario, se clasifica como un rectángulo.
- **Círculo:** Si la figura tiene más de cuatro vértices, se analiza si puede ser un círculo. Esto se hace comparando el área del contorno con el área de un círculo que lo podría encerrar. Si la diferencia entre estas áreas es pequeña, se clasifica como un círculo.

Al adaptar la detección de formas a variaciones en tamaño y orientación, el sistema es capaz de reconocer y validar las paletas mostradas por los niños incluso cuando estas se encuentran a distintas distancias y con cualquier inclinación respecto de los ejes de la imagen.

En resumen, la detección de formas convierte los bordes detectados en una representación donde las formas de interés están claramente identificadas, lo que facilita la evaluación precisa de sus respuestas (*Figura 7*).



*Figura 7. Detección de forma*

### III. RESULTADOS

En primera instancia se probaron individualmente cada una de las etapas del procesamiento con el objetivo de ajustar los parámetros para su uso real. Una vez que se verificó el funcionamiento correcto de cada etapa, se procedió a realizar pruebas generales del algoritmo en diversas condiciones de iluminación, proximidad a la cámara, tipos de cámara y escenas con objetos en el fondo.

En general, el algoritmo arrojó resultados satisfactorios en la mayoría de los casos, detectando la paleta correcta en fracciones de segundo. La elección de técnicas de procesamiento permite que el detector funcione en tiempo real, proporcionando al usuario una experiencia fluida y sin interrupciones.

Sin embargo, en situaciones muy particulares donde la cámara capta objetos grandes con un tono similar al de la paleta buscada, pueden producirse falsos positivos. Esto puede ocurrir, por ejemplo, si el niño lleva una prenda de un color liso similar al de la paleta buscada y en algún momento

aparece un área que se asemeje a la forma deseada. Aunque esta situación es infrecuente, se podría considerar la adición de un temporizador en el algoritmo que verifique que la figura buscada se mantenga en pantalla durante varios cuadros consecutivos, de modo que las áreas que aparecen casualmente sean ignoradas.

Asimismo, durante las pruebas se comprobó la importancia de la iluminación, ya que la percepción del color de la paleta depende significativamente de esta. En condiciones de iluminación muy deficiente, la detección puede fallar, por lo que se recomienda jugar en un ambiente mínimamente iluminado, ya sea con luz natural o artificial. Si bien se podría ajustar los parámetros del rango de intensidad para admitir colores con muy baja intensidad, esta medida no es necesaria si se sigue la recomendación de mantener una iluminación adecuada.

#### IV. CONCLUSIONES

En resumen, el exergame desarrollado representa una herramienta prometedora para la intervención terapéutica en niños con TEA, combinando elementos de juego interactivo y tecnología de procesamiento de imágenes. Su accesibilidad y bajo costo lo hacen viable para un amplio rango de usuarios, ofreciendo una alternativa atractiva y efectiva a las terapias tradicionales. Este enfoque innovador no solo fomenta la actividad física y el desarrollo motor, sino que también estimula las habilidades cognitivas a través de la identificación de formas y colores, mejorando así la experiencia de aprendizaje y la interacción de los niños con su entorno.

A pesar de los resultados generalmente satisfactorios, se identificaron algunas limitaciones, como la susceptibilidad a falsos positivos en condiciones específicas y la dependencia de una iluminación adecuada. Estas limitaciones subrayan la importancia de continuar afinando el algoritmo, incorporando, por ejemplo, un temporizador que verifique la persistencia de las figuras detectadas para reducir los falsos positivos y calibración automática que permita ajustar los límites de color en tiempo real, adaptándose a diferentes condiciones de iluminación.

Además, futuras iteraciones del juego podrían considerar la incorporación de más colores y formas, así como la posibilidad de personalizar los desafíos según las necesidades específicas de cada niño, aumentando así la flexibilidad y efectividad de la herramienta. Con estos ajustes adicionales, el sistema puede optimizarse para diversas condiciones de uso, asegurando una experiencia más robusta y fiable.

#### V. REFERENCIAS

- [1] Lord, C., & Bishop, S. L. (2010). Autism spectrum disorders: Diagnosis, prevalence, and services for children and families. Social Policy Report, 24(2).
- [2] American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.
- [3] Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36(1).
- [4] Anzalone, S. M., Tkach, J. A., & Harrell, B. L. (2014). Exergaming and children with autism spectrum disorder: The future of health promotion. *Games for Health Journal*, 3(2), 87-90.
- [5] Parsons, S., & Cobb, S. (2011). State-of-the-art of virtual reality technologies for children on the autism spectrum. *European Journal of Special Needs Education*, 26(3), 355-366.
- [6] K. Caro, M. Tentori, A. I. Martínez-García, and I. Zavala-Ibarra, "FroggyBobby: An exergame to support children with motor problems practicing motor coordination exercises during therapeutic interventions," *Comput. Human Behav.*, vol. 71, pp. 479-498, Jun. 2017.
- [7] F. L. Cibrian, O. Peña, D. Ortega, and M. Tentori, "BendableSound: An elastic multisensory surface using touch-based interactions to assist children with severe autism during music therapy," *Int.*
- [8] Monarca Pintle, G. I., Morales Téllez, A., Cibrian Robles, F. L., & Tentori Espinosa, M. E. (2018). PaU: Diseño de un exergame en una superficie elástica para niños con autismo severo.
- [9] Chen, W. (2012). Multimodal-based computer-supported collaborative learning for children with autism spectrum disorders. *Universal Access in the Information Society*, 11(4), 321-332.
- [10] Suzuki, S., & Abe, K. (1985). "Topological Structural Analysis of Digitized Binary Images by Border Following." *Computer Vision, Graphics, and Image Processing*, 30(1), 32-46