Literature Review

Interactive Educational Puzzle Web Platform Utilizing Computer Vision

Assignment 3 - Or Reshef S 324064849 Gal Azoulay 323859967

Introduction

The integration of Computer Vision (CV) technology into educational environments represents a paradigm shift in how children interact with digital learning tools. Traditional digital games validate user input through coordinate-based checking, determining success by whether a child clicked or placed an item in the predetermined correct location. Modern CV algorithms, however, enable real-time visual assessment that mimics human perception, evaluating whether a selection demonstrates genuine visual understanding rather than mere mechanical placement.

This literature review examines current research across three critical domains: the impact of digital puzzle games on cognitive development in early childhood education, the application of computer vision techniques in puzzle-solving systems and the technical foundations required for visual validation. By synthesizing these areas, this review identifies a significant gap in the current landscape: existing educational games validate answers through coordinate checking rather than visual analysis, while existing CV research focuses on robotic automation rather than educational assessment for children. The proposed platform aims to bridge this gap by employing CV as the primary validation mechanism, creating an endless dynamic puzzle experience that validates children's selections based on actual visual coherence, like edge continuity, color compatibility and texture consistency, rather than predetermined answer keys stored in databases.

Competitor Analysis

Several online educational puzzle platforms currently serve the children's learning market, but all rely on coordinate-based validation rather than visual analysis. The website *Digipuzzle.net* offers comprehensive educational puzzle games including jigsaw puzzles, memory games and logic challenges. However, the platform provides a finite, pre-selected collection of puzzles determined entirely by the developers, with no mechanism for users to request specific themes or topics. Validation occurs through simple drag-and-drop mechanics that check only whether a piece was placed in the designated target zone. The system cannot distinguish between a correct piece and an incorrect piece placed in the same location. Additionally, pieces that are incorrectly positioned simply fail to "stick" to the board, allowing children to repeatedly attempt placement in different locations until finding the correct coordinates through trial-and-error rather than visual reasoning.

Another website named *RoomRecess.com* provides jigsaw puzzles for young children with bright colors and simple themes. While the platform offers two difficulty levels (Simple and Hard), both levels present traditional jigsaw puzzles where all pieces are visible simultaneously and must be assembled in their entirety. The content library is predetermined by developers, offering no personalization based on children's interests. Validation remains entirely position-dependent: pieces only attach to the board when dropped in their exact predetermined locations, encouraging a brute-force approach of testing multiple positions until one succeeds. Similarly, *HappyClicks.net* and *SafeKidGames.com* websites employ identical mechanisms and limitations, except that they lack any difficulty level differentiation whatsoever.

All reviewed platforms share fundamental limitations. First, content is finite and predetermined. Once children complete available puzzles, the learning experience ends. Second, children cannot select puzzle themes or topics matching their interests. All content is chosen by developers. Third, validation occurs through coordinate checking rather than visual analysis, enabling children to "solve" puzzles through systematic trial-and-error placement without genuine visual understanding. Fourth, feedback is binary (piece attaches or doesn't) rather than visually grounded. Fifth, traditional jigsaw formats require assembling all pieces simultaneously, whereas image completion formats that progressively reveal missing sections are unavailable.

Our proposed platform addresses these limitations comprehensively. It employs computer vision validation that analyzes visual coherence (edge continuity, color compatibility, texture consistency) rather than coordinates, preventing trial-and-error solving. It provides infinite puzzle variations through dynamic content generation via Unsplash API, ensuring inexhaustible educational content. It enables children to select puzzle themes through category selection or keyword input, making content personally relevant. It implements progressive difficulty across scalable levels of complexity, transitioning from simple image completion at beginner stages to complex multi-piece puzzles at advanced levels. Most critically, pieces are validated based on visual correctness regardless of placement location, requiring genuine visual reasoning rather than mechanical coordinate discovery.

Digital Game-Based Learning and Cognitive Development

Recent systematic reviews establish robust evidence for game-based learning effectiveness in promoting cognitive, social and emotional development in young children. Alotaibi (2024) conducted a comprehensive meta-analysis examining 136 studies (from 2013-2023) involving 1,426 participants aged 3-8 years. The meta-analysis measured effect sizes using Hedges' g, where larger values indicate stronger impacts (0.2=small, 0.5=medium, 0.8=large effect size). The analysis revealed significant effects across developmental domains: cognitive development (g=0.46), social development (g=0.38), emotional development (g=0.35), motivation (g=0.40) and engagement (g=0.44). Critically, puzzle games demonstrate significantly larger cognitive effects (g=0.63) compared to other game types (g=0.31), underscoring puzzle-based learning as particularly effective for enhancing children's cognitive abilities.

Wang et al. (2023) developed "Tree of Wisdom: Color Matching and Touch", a multiplayer digital puzzle system for children aged 4-6 combining physical controllers with computer vision technology that provides immediate, multi-sensory feedback via an interactive projection table. Visual animations and auditory cues respond to every action, error sounds guide incorrect attempts while success animations reinforce correct choices. Evaluation with 58 children demonstrated that this positive, immediate feedback mechanism effectively promotes cognitive development and motor skills by allowing children to learn from both successful and unsuccessful attempts without penalty.

Our project builds directly upon these established principles, recognizing puzzle-based games as optimal for cognitive development and immediate encouraging feedback as essential for effective learning. By implementing CV-based validation that assesses visual coherence (edge continuity, color compatibility, texture consistency) and visual analysis of contextual relationships for advancing validation.

Computer Vision in Puzzle-Solving Systems

The technical feasibility of implementing complex visual processing on accessible devices has been confirmed by recent developments. Sigut et al. (2020) demonstrated OpenCV's effectiveness in mobile applications designed to teach computer vision concepts to students, confirming that modern mobile devices possess sufficient computational power to execute real-time image processing algorithms such as edge detection, feature matching, and color analysis without significant latency. This validates the technical viability of web-based or mobile applications performing instant visual validation on users' devices without requiring server-side processing.

The specific application of computer vision to puzzle-solving has been extensively researched in robotic contexts, providing valuable methodological insights for educational validation systems. Ma et al. (2023) developed a vision-based system for robotic jigsaw puzzle solving achieving an 87.1% success rate on puzzles ranging from 35 to 70 pieces. Their algorithm employed two distinct approaches depending on data availability. When the original image was available, they used SIFT (Scale-Invariant Feature Transform) for feature extraction combined with RANSAC (Random Sample Consensus) for matching validation. In scenarios without the original image, they utilized a reconstruction algorithm based on edge compatibility, applying Sobel filter similarity to calculate the gradients at puzzle boundaries. This was implemented through a greedy "square expansion" strategy that identified a "Best Piece" using Hausdorff distance concepts and progressively built the puzzle outward

Critically, Ma et al.'s approach focused on instructing a robot to physically manipulate and assemble puzzle pieces, whereas the proposed platform adapts these validation algorithms for human-computer interaction. Rather than the system solving the puzzle, the system validates whether the child's selection demonstrates appropriate visual understanding, a fundamental shift from automation to assessment that preserves educational value while providing immediate, visually grounded feedback.

Image Segmentation and Regional Manipulation

Modern computer vision relies on sophisticated techniques for understanding and manipulating image regions. Minaee et al. (2021) conducted a comprehensive survey of image segmentation using deep learning, examining architectures including U-Net and encoder-decoder models that enable intelligent partitioning of images into meaningful regions. These segmentation techniques provide the theoretical foundation for our project to divide images along natural semantic boundaries rather than arbitrary grids, ensuring puzzle pieces represent coherent visual elements. By leveraging these deep learning foundations, our project maintains the flexibility to validate selections based not only on traditional visual features (edges, colors, textures) but also on high-level semantic coherence, verifying that a specific object part aligns logically with the surrounding visual context in future iterations.

Complementing segmentation, recent research on masked image modeling offers methodological context for the efficient programmatic manipulation of these regions. Sarkar et al. (2025) demonstrated that region masking techniques (originally explored for video object detection) can identify and process specific image areas while maintaining computational efficiency. These findings inform our project's strategy of systematically occluding selected regions to create visual challenges that demand genuine pattern recognition rather than simple memorization.

Algorithmic Foundations: Edge Detection and Color Analysis

The validation of puzzle piece correctness relies fundamentally on analyzing visual continuity at boundaries. Edge detection algorithms provide the technical foundation for this analysis. Canny edge detection remains one of the most widely used methods for identifying boundaries in images, employing multi-stage processing including Gaussian filtering for noise reduction, gradient magnitude calculation, non-maximum suppression, and double thresholding with hysteresis. The Sobel operator, a simpler first-order derivative method, computes image gradients using 3x3 convolution kernels, providing computationally efficient edge detection particularly suitable for real-time applications.

Comparative studies of edge detection algorithms demonstrate that Canny consistently achieves superior detection accuracy with high sensitivity and robustness against noise, while Sobel excels in computational efficiency and real-time processing scenarios. For educational puzzle validation, edge continuity across puzzle piece boundaries serves as a primary indicator of correct placement, making edge detection algorithms essential to the validation pipeline.

Color compatibility analysis provides another crucial validation dimension. The HSV (Hue, Saturation, Value) color space, which separates chromatic information from intensity, proves particularly effective for color-based image analysis. Unlike RGB color space, where color and brightness information are entangled, HSV allows independent manipulation and comparison of color properties. Histogram-based color analysis in HSV space enables quantitative assessment of color similarity between adjacent image regions, supporting validation of whether selected puzzle pieces exhibit appropriate color continuity with surrounding image context.

Dynamic Content Generation and Research Gap

Our project leverages the Unsplash API to provide infinite, dynamically generated puzzle content. Unsplash offers access to over 6.5 million high-quality photographs contributed by 350,000 photographers, with the platform receiving over 250 million image searches monthly. By integrating this vast, continuously expanding repository, our project overcomes the limitation of fixed content libraries found in traditional games, ensuring that every session presents fresh visual challenges that sustain long-term engagement.

This literature review reveals a significant research gap: existing educational games validate puzzle solutions through coordinate-based checking rather than visual analysis. Conversely, existing computer vision research on puzzle-solving, exemplified by Ma et al. (2023), focuses on robotic automation rather than educational contexts where the goal is assessing human visual reasoning and providing meaningful feedback supporting learning.

Conclusion

This literature review synthesizes research across game-based learning, computer vision puzzle-solving, and image processing techniques, revealing both the established benefits of puzzle-based education and the technical feasibility of sophisticated visual validation systems. The reviewed studies confirm that puzzle games significantly enhance cognitive development (Alotaibi, 2024, Wang et al., 2023), that CV algorithms can achieve high-accuracy puzzle validation (Ma et al., 2023), that these algorithms run efficiently on accessible devices (Sigut et al., 2020), and that modern image processing techniques enable intelligent image manipulation (Minaee et al., 2021, Sarkar et al., 2025).

The critical gap identified is that existing educational games validate through coordinate-checking rather than visual analysis, while existing CV research focuses on automation rather than educational assessment. Our project addresses this gap by employing CV as the primary validation mechanism, creating an environment where children receive immediate, visually grounded feedback based on actual image analysis rather than predetermined answer keys. Combined with dynamic content generation via Unsplash API, this approach creates an endless, adaptive puzzle experience maintaining engagement while supporting sustained cognitive development.

References

Alotaibi, M. S. (2024). *Game-based learning in early childhood education: A systematic review and meta-analysis.* Frontiers in Psychology, 15, 1307881.
https://doi.org/10.3389/fpsyg.2024.1307881

Ma, C.-H., Lu, C.-L., & Shih, H.-C. (2023). *Vision-based jigsaw puzzle solving with a robotic arm. Sensors*, *23*(15), 6913.
https://doi.org/10.3390/s23156913

Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2021*). Image segmentation using deep learning: A survey.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(7), 3523–3542.
https://doi.org/10.1109/TPAMI.2021.3059968

Sarkar, S., Datta, G., Kundu, S., Zheng, K., Bhattacharyya, C., Beerel, P. A. (2025). *MaskVD: Region masking for efficient video object detection.* IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 1955-1964.
https://doi.org/10.1109/WACV61041.2025.00197

Sigut, J., Castro, M., Arnay, R., & Sigut, M. (2020). *OpenCV basics: A mobile application to support the teaching of computer vision concepts*. IEEE Transactions on Education, 63(4), 328–335.
https://doi.org/10.1109/TE.2020.2993013

Unsplash. (2024). *Unsplash API Documentation*.
https://unsplash.com/documentation

Wang, C.-M., Lee, B.-T., & Lo, T.-Y. (2023). *The design of a novel digital puzzle gaming system for young children's learning by interactive multi-sensing and tangible user interfacing techniques*. Sustainability, 15(4), 3036.
https://doi.org/10.3390/su15043036