

Uso de Game Analytics no Jogo Pac-Man para Análise de Jogadores: Uma Revisão Sistemática da Literatura

Guilherme Araujo Lira de Menezes¹, Luciana Rita Guedes¹, Marcelo da Silva Hounsell¹

¹ Departamento de Ciência da Computação
Universidade do Estado de Santa Catarina (UDESC) – Joinville, SC – Brasil

guilherme.menezes@edu.udesc.br;
{luciana.guedes,marcelo.hounsell}@udesc.br

Abstract. *This article presents a systematic literature review on the use of the Pac-Man game in scientific research focused on data analysis. A total of 25 studies were selected from 7 academic databases, using both inclusion and exclusion criteria. The review identified the data collection methods employed (such as automated collection, screen capture, and sensors) and the different types of data extracted, organized into game characters, game environment, and external sources. The results reveal a wide variety in the use of Pac-Man as a study environment but also highlight a lack of standardization in data description and the still limited application of systematic analyses among internal game variables.*

Resumo. *Este artigo apresenta uma revisão sistemática da literatura sobre o uso do jogo Pac-Man em pesquisas científicas voltadas à análise de dados. Foram selecionados 25 estudos a partir de 7 bases acadêmicas, com critérios de inclusão e de exclusão. A revisão identificou os métodos de coleta empregados (como coleta automática, captura de tela e sensores) e os diferentes tipos de dados extraídos, organizados entre personagens, ambiente do jogo e fontes externas. Os resultados revelam ampla diversidade no uso do Pac-Man como ambiente de estudo, mas também destacam a falta de padronização na descrição dos dados e a aplicação ainda limitada de análises sistemáticas entre variáveis internas do jogo.*

1. Introdução

O uso de jogos digitais em pesquisas acadêmicas destaca-se pela capacidade de simular ambientes controlados e interativos que permitem a análise da tomada de decisão, a adaptação a cenários dinâmicos, a resolução de problemas e a reflexão sobre consequências. Segundo Gee (2003), os jogos oferecem um espaço único para experimentação, no qual os jogadores assumem papéis, tomam decisões e enfrentam resultados em um contexto seguro e imersivo. Essa característica os torna ferramentas relevantes para áreas como a psicologia cognitiva, ao possibilitar o estudo de processos de aprendizado e resolução de problemas (Green; Bavelier, 2012). Além disso, a coleta de dados pode ocorrer de maneira não invasiva, permitindo obter informações diretamente dos usuários (Mahlmann et al., 2010).

Pesquisas recentes indicam que, quando bem projetados e contextualizados, os jogos digitais também podem reproduzir dinâmicas sociais e psicológicas de forma fidedigna. Mecânicas como recompensas e *feedbacks* imediatos já foram empregadas em

estudos para investigar motivações e padrões de comportamento (Deterding, 2015), bem como interações consistentes com teorias da psicologia social (Johnson et al., 2004). Nesse cenário, a área de game analytics consolidou-se como campo voltado à análise sistemática de dados de design, uso e impacto dos jogos. De acordo com El-Nasr et al. (2013), game analytics trata da aplicação de técnicas de análise para apoiar decisões em diferentes níveis, abrangendo desde aspectos de desenvolvimento até pesquisas com usuários.

O jogo Pac-Man, lançado em 1980 pela Namco, é um dos títulos mais influentes da história dos jogos digitais e permanece como ferramenta recorrente em pesquisas acadêmicas. Sua mecânica simples, baseada em coletar pontos e evitar fantasmas, tornou-o apropriado para estudos de inteligência artificial e aprendizado de máquina (Santos Junior et al., 2019). Além disso, sua versatilidade e longa trajetória já foram associadas a investigações em áreas como psicologia, biologia e inteligência computacional (Rohlfshagen et al., 2018).

Apesar desse uso consolidado, observa-se que no caso do Pac-Man, análises quantitativas de elementos internos, como consumo de pastilhas, padrões de movimento e interações com fantasmas, ainda são pouco exploradas. Na literatura sobre jogos digitais em geral são mais comuns estudos que priorizam variáveis externas, como motivações dos jogadores (Yee, 2006) e fatores sociais. Essa lacuna limita a compreensão das dinâmicas entre os elementos jogo e de como influenciam o desempenho do jogador.

Diante desse cenário, este trabalho apresenta uma revisão sistemática da literatura sobre o uso do Pac-Man em pesquisas científicas, com ênfase nos métodos de coleta de dados e nas abordagens de análise empregadas. O objetivo é organizar e sintetizar as práticas relatadas nos estudos, abrangendo tanto os dados dos personagens quanto do próprio jogo, bem como dados externos e ainda os métodos utilizados em sua análise. Espera-se que os resultados ofereçam uma visão consolidada sobre como o Pac-Man tem sido empregado como ambiente de investigação, evidenciando seu potencial como recurso metodológico para estudos futuros que adotem esse jogo como objeto de análise.

2. Metodologia da Revisão Sistemática

A condução desta revisão foi fundamentada nas diretrizes propostas por (Kitchenham; Charters, 2007), conforme segue nesta seção.

2.1. Objetivos da Revisão

Para realização da revisão foram definidas as Questões de Pesquisa (QP) do estudo. A questão principal é: “Como o jogo Pac-Man tem sido analisado em pesquisas científicas?”. Além da questão principal, para direcionar os dados coletados, foram formuladas as seguintes questões secundárias:

- Q1: Quais métodos têm sido utilizados para a coleta e o tratamento dos dados?
- Q2: Quais dados são colhidos para serem utilizados nas análises?
- Q3: Quais abordagens ou técnicas de análise são aplicadas aos dados coletados?

2.2. Seleção de Fontes de Busca

A escolha das fontes de busca foi fundamentada no estudo de Buchinger, Cavalcanti e Hounsell (2014), onde foram selecionadas plataformas com melhor avaliação,

complementadas pelo Periódico CAPES que é amplamente utilizado no contexto nacional. Os mecanismos de busca selecionados como fontes para esta pesquisa foram: Periódico CAPES, ACM Digital Library, ScienceDirect, IEEE Xplore, Springer Link, Scopus e Web of Science.

2.3. Estratégias de Busca

A estratégia de busca foi desenvolvida de forma empírica, por meio de testes iterativos nas plataformas selecionadas, visando à identificação de estudos relevantes para a pesquisa. Nesse processo, foram considerados três critérios principais para a escolha dos termos e a construção da frase de busca. O primeiro foi o foco específico no termo “Pac-Man”, utilizado para restringir a busca a estudos centrados no jogo e evitar a dispersão causada por termos genéricos como “jogos” (ou “*games*”); o segundo critério foi a escolha de expressões da área de análise de dados em jogos, e o terceiro a inclusão de termos identificados em pesquisas preliminares, como “*Gameplay Analysis*” e “*Player Behavior*”, com o objetivo de ampliar a abrangência da busca e captar estudos relevantes que empregam essas nomenclaturas.

Tabela 1. Frases de Busca em Inglês e em Português

Idioma	Frase de Busca Correspondente
Inglês	("Pac-Man" OR "PacMan") AND ("Data Science" OR "Game Analytics" OR "Game Metrics" OR "Data Mining" OR "Gameplay Analysis" OR "Player Behavior")
Português	("Pac-Man" OR "PacMan") AND ("Ciência de Dados" OR "Game Analytics" OR "Métricas de Jogos" OR "Mineração de Dados" OR "Análise de Jogabilidade" OR "Comportamento do Jogador")

Fonte: Elaborado pelos autores (2025)

A partir desses critérios, a frase final construída apresentada na Tabela 1. A versão em português foi utilizada exclusivamente na base do Periódico CAPES, enquanto a versão em inglês foi aplicada nas demais plataformas de busca. Cada frase de busca foi aplicada ao Título, Resumo e Palavras-Chave nos mecanismos de busca já identificados.

2.4. Critérios de Inclusão e de Exclusão

Os critérios de seleção utilizados estão detalhados na Tabela 2, com ênfase em garantir a relevância dos trabalhos encontrados. O recorte temporal de 10 anos (2015–2025) foi definido como uma decisão metodológica, com o objetivo de concentrar a análise em estudos recentes e alinhados às tecnologias mais atuais.

Tabela 2. Listas de Critérios de Seleção dos Estudos

Critérios de Inclusão (CI)		Critérios de Exclusão (CE)	
#	Descrição	#	Descrição
1	Artigos disponíveis gratuitamente ou acessíveis via Portal da CAPES	1	Artigos que não descrevem coleta e/ou análise de dados
2	Artigos publicados nos últimos 10 anos (2015 – 2025)	2	Artigos não relacionados à análise de dados
3	Artigos em inglês ou português	3	Artigos duplicados
4	Mais do que 5 páginas	4	Artigos que não contemplam o Pac-Man
5	Estudo Primário (não incluir estudos de revisão)		

Fonte: Elaborado pelos autores (2025)

2.5. Resultados de Busca e Seleção

A Tabela 3 apresenta a quantidade de artigos retornados por cada fonte e o número final de artigos selecionados após aplicação dos critérios.

Tabela 3. Quantitativo de Artigos Identificados e Selecionados

Fonte de Busca	Artigos Encontrados	Após Critérios de Inclusão (C.I.)	Artigos Selecionados Após Critérios de Exclusão (C.E.)
ACM Digital Library	127	12	1
ScienceDirect	80	46	1
IEEE Xplore	26	10	6
Springer Link	554	45	1
Scopus	12	3	2
Web of Science	5	3	0
Periódico CAPES	428	114	14
TOTAIS	1.232	233	25

Fonte: Elaborado pelos autores (2025)

Esses 25 estudos compõem o *corpus* para responder tanto à questão principal quanto às questões secundárias propostas no objetivo desta revisão.

2.6. Trabalhos Selecionados

Na Tabela 4 são apresentados os resultados com base nas questões secundárias da pesquisa. Os artigos selecionados foram identificados com a letra “A” e um número único para facilitar sua referência ao longo deste trabalho.

3. Análise dos Resultados

Nesta seção serão apresentados os resultados da análise dos métodos utilizados para a coleta e o tratamento dos dados nos estudos selecionados, com foco nos dados extraídos do próprio jogo Pac-Man. A análise procura considerar tanto a frequência com que diferentes abordagens são utilizadas quanto às particularidades dessas abordagens.

Tabela 4. Artigos Selecionados para a Revisão Sistemática

ID	Autor(es)	ID	Autor(es)
A01	(An; Du, 2022)	A14	(Neufeld et al., 2022)
A02	(Banerjee et al., 2019)	A15	(Safak; Bostanci; Soylicicek, 2016)
A03	(Blom; Bakkes; Spronck, 2019)	A16	(Junior et al., 2019)
A04	(Butt et al., 2023)	A17	(Schrum; Miikkulainen, 2016)
A05	(Cai et al., 2023)	A18	(Silva; Júnior, 2020)
A06	(Costa; Giancarlo; Adamatti, 2022)	A19	(Sledge; Principe, 2017)
A07	(Cowley; Charles, 2016)	A20	(Tamassia et al., 2018)
A08	(Fachantidis; Taylor; Vlahavas, 2017)	A21	(Wang et al., 2020)
A09	(Foderaro; Swinger; Ferrari, 2017)	A22	(Wolterink; Bakkes, 2021)
A10	(Hatschka; Ciabattini; Eiter, 2023)	A23	(Yang et al., 2022)
A11	(Maddipati et al., 2020)	A24	(Zhang et al., 2018)
A12	(Miranda; Sánchez-Ruiz; Peinado, 2021)	A25	(Zikky, 2016)
A13	(Morosan; Poli, 2017)		

Fonte: Elaborado pelos autores (2025)

3.1. Coleta de Dados

Dos 25 estudos analisados, 22 descrevem o processo de coleta de dados, sendo a maioria realizada automaticamente durante a execução. A primeira parte desta seção apresenta a distribuição dos tipos de agente utilizados e caracteriza a dosimetria das coletas humanas. A Tabela 5 resume os tipos de agentes empregados nos estudos.

Tabela 5. Distribuição dos Tipos de Agente nos Artigos Selecionados

Agente	Frequência	Artigos (Índice)
Autônomo	17	A01; A05; A06; A08; A09; A10; A11; A13; A14; A15; A16; A17; A18; A19; A20; A24; A25
Humano	5	A02; A03; A04; A07; A22
Misto	2	A12; A20
Animal	2	A21; A23

Fonte: Elaborado pelos autores (2025)

A maior parte dos estudos (17 deles) utilizou agentes computacionais para jogar Pac-Man. Esses trabalhos priorizam o desenvolvimento e avaliação de algoritmos de tomada de decisão, aprendizado de máquina e inteligência artificial, sem envolver diretamente participantes humanos.

A participação exclusiva de jogadores humanos ocorreu em 5 estudos. No A02, 10 indivíduos jogaram 15 sessões de 30 segundos cada, realizando em seguida autoavaliações da experiência. O A03 contou com 32 participantes, que jogaram uma vez em 3 diferentes dificuldades do Pac-Man, enquanto suas expressões faciais eram registradas automaticamente. O A04 utilizou registros de jogadas humanas para treinar agentes imitativos, embora sem detalhar número de participantes ou volume de dados. No A07, 100 indivíduos jogaram livremente para familiarização, sendo considerada para análise a partida com maior número de níveis concluídos após a terceira tentativa. Já o A22 envolveu 135 participantes, que jogaram cinco fases do jogo e responderam questionários de frustração a cada 30 segundos.

Apenas 2 estudos combinaram agentes humanos e computacionais. No A12, 5 designers profissionais alternam o controle do Pac-Man com um agente autônomo, realizando intervenções mais frequentes no início e pontuais nas etapas posteriores. O A20 registrou partidas de 2 agentes automáticos (1000 cada) e 2 jogadores humanos (30 cada), possibilitando comparação entre ambos os tipos de agentes.

Em 2 trabalhos, o Pac-Man foi jogado por macacos. No A21, foram gravados vídeos de 1297 partidas completas realizadas por macacos treinados, resultando em aproximadamente 1,2 milhão de amostras de movimentos. No A23, a coleta incluiu 74 sessões de teste após o treinamento, realizadas quando o desempenho dos animais atingiu nível considerado satisfatório.

A Tabela 6 apresenta a classificação dos métodos de coleta de dados adotados nos estudos analisados. Entre os estudos que utilizaram captura de tela ou extração de imagens como método de coleta, houve algumas diferenças nas particularidades do método. Em A01, a captura foi descrita com aquisição de quatro *frames* a cada 0,16 segundos diretamente da execução do jogo na plataforma de jogos *Steam*. Nos estudos A05, A09 e A19, capturas de tela foram realizadas, mas sem especificação de frequência ou sistematização. Os estudos A20 e A21 se diferenciam pois realizaram a extração de

imagens a partir de gravações de sessões humanas ou vídeos de *gameplay* de macacos, respectivamente, e não diretamente da execução do jogo em tempo real.

Tabela 6. Distribuição dos Tipos de Coleta de Dados nos Artigos Selecionados

Coleta	Frequência	Artigos (Índice)
Coleta automática não especificada	9	A08; A12; A13; A16; A17; A18; A22; A24; A25
Captura de tela	6	A01; A05; A09; A19; A20; A21
Coleta automática via módulo / <i>logger</i>	4	A04; A07; A10; A14
Questionário	4	A02; A03; A12; A22
Equipamento externo	4	A02; A03; A12; A21
Não se aplica	3	A06; A11; A15

Fonte: Elaborado pelos autores (2025)

Alguns estudos utilizaram equipamentos externos para a coleta de informações fisiológicas e comportamentais. Entre eles, o A02 realizou registros de sinais de respiração e atividade cardíaca, o A03 e o A12 utilizaram câmeras para capturar expressões faciais durante a partida e o A21 monitorou movimentos oculares e dilatação pupilar.

Entre os estudos que aplicaram questionários, todos utilizaram escalas do tipo Likert para a obtenção de dados subjetivos. No A02, os participantes se autoavaliaram usando o *Self-Assessment Manikin* (SAM) em uma escala de 9 pontos. O A03 aplicou um questionário de avaliação de dificuldade percebida, utilizando uma escala de 7 pontos. No A12, foi aplicado um questionário com perguntas fechadas em escala de 5 pontos, seguido por espaço para comentários abertos. Já o A22 utilizou um questionário para avaliação contínua do nível de frustração geral, também em escala de 5 pontos.

Por fim, dos estudos analisados, alguns indicaram a realização de coleta automática de dados durante a execução do jogo, mas sem descrever de forma explícita o procedimento de armazenamento ou registro das informações geradas. Nesses casos, infere-se apenas que a captura de dados ocorre concomitantemente ao ambiente de jogo, sem detalhamento sobre a estrutura dos arquivos, frequência de amostragem ou método.

3.2. Dados Utilizados

Os estudos demonstram ampla variedade na seleção dos dados extraídos do jogo Pac-Man, os quais foram organizados em duas categorias principais: dados dos agentes (Pac-Man e fantasmas); e dados gerais do ambiente. Além desses, uma parte dos estudos também incorporou dados externos, como expressões faciais ou sinais fisiológicos, que são tratados separadamente ao final desta seção.

3.2.1 Dados dos Personagens (Pac-Man e Fantasmas)

Os dados relacionados aos personagens do jogo - Pac-Man e fantasmas - são amplamente empregados nos estudos analisados, aparecendo em 21 dos 25 artigos selecionados e compondo o principal núcleo das representações de estado utilizadas nas análises. Esses dados descrevem o comportamento dos personagens dentro do labirinto e são centrais para a maioria dos modelos de tomada de decisão. A Tabela 7 apresenta a relação completa das informações extraídas nos estudos.

Entre os elementos mais recorrentes estão a posição e as quatro direções de movimento utilizadas para representar o estado espacial dos personagens no labirinto e possibilitar o cálculo de trajetórias, distâncias e previsões de movimento. A distância foi representada de formas distintas nos estudos analisados: a distância em número de passos foi utilizada em A20, A23 e A19, representada como o número de células percorridas no labirinto, associada ao menor caminho entre personagens ou entre personagens e elementos do ambiente.

Tabela 7. Distribuição de Dados dos Personagens nos Artigos Selecionados

#	Tipo de Dados	Origem	Freq.	Artigos (Índice)
Dados Espaciais				
1	Posição	Original	16	A01; A04; A07; A08; A09; A10; A11; A14; A16; A17; A18; A19; A21; A22; A23; A25
2	Direção	Original	10	A09; A11; A12; A17; A18; A19; A20; A21; A23; A24
3	Distância	Derivado	12	A07; A08; A12; A13; A16; A17; A18; A19; A20; A21; A23; A25
4	Velocidade	Derivado	2	A09; A22
5	Mesmo túnel que outro personagem	Derivado	1	A19
6	Ação executada	Original	6	A01; A10; A11; A12; A17; A20
7	Ações válidas	Derivado	3	A07; A17; A22
Dados Cumulativos				
8	Número de Elementos Comidos (Pastilhas, Frutas, Fantasmas)	Derivado	4	A07; A16; A17; A22
9	Número de Mortes	Derivado	2	A07; A22
10	Número de Teletransportes	Derivado	2	A07; A22
Dados Temporais				
11	Tempo Parado	Derivado	1	A22
12	Instante da Morte	Original	1	A22
13	Tempo para Reviver	Original	1	A16
14	Duração do Estado (Tempo restante; Tempo total)	Derivado	4	A08; A12; A16; A22
Dados Situacionais				
15	Estado do Personagem	Original	14	A07; A08; A09; A10; A13; A14; A16; A17; A18; A19; A20; A21; A22; A23
16	Existência de Nó de Junção	Derivado	1	A17
Dados Identificadores				
17	Tipo de Fantasma (Cor; Nome)	Original	2	A16; A23
Dados de Desempenho				
18	Total de Passos	Derivado	1	A06
19	Pontuação Estimada por Ação	Derivado	1	A11
Não se Aplica				
20	Não se Aplica	-	4	A02; A03; A05; A15

Fonte: Elaborado pelos autores (2025)

Em A18, foram representadas três métricas: distância euclidiana, distância de Manhattan e número de passos. A distância euclidiana também foi utilizada em A12. Nos casos de A07 e A08, embora não haja especificação explícita da métrica adotada, as

descrições sugerem o uso de distâncias discretas, baseadas no número de passos.

Outros dados relativos à movimentação do Pac-Man são a ação executada - o movimento realizado - e as ações válidas que indicam as possibilidades de movimento permitidas a partir de cada estado registrado, informação que depende diretamente da estrutura do ambiente. Também foi computado, em um dos estudos, se o Pac-Man e um fantasma estavam no mesmo túnel, um dado espacial derivado de suas posições relativas.

Outro elemento considerado inclui a velocidade dos personagens, extraída como derivada das posições em determinados intervalos de tempo.

Entre os dados cumulativos, aparecem contagens como o número de elementos comidos pelo Pac-Man - pastilhas, pastilhas de poder, frutas e fantasmas - além do número de mortes e teletransportes realizados. No caso dos fantasmas, em A22 também foram registradas o número de mortes causadas por cada um deles.

Os dados temporais incluem o tempo em que o Pac-Man ficou parado, os instantes das suas mortes e o tempo que o fantasma levou para reviver após captura. Para os fantasmas também foi registrada a duração em um estado modificado.

Nos dados situacionais está o estado dos fantasmas (como “normal”, “assustado”, “comestível”). Os trabalhos A07, A09 e A22 utilizam representações multicategóricas, diferenciando múltiplos modos de comportamento dos fantasmas, como “Perseguir”, “Espalhar” (recuar para os cantos do mapa), “Assustado” (fantasma vulnerável após o consumo de pastilha de poder), “Comido” (retornando à base após ser capturado) e “Esperando” (fantasma parado na base). Esses trabalhos caracterizam de forma mais refinada as condições dinâmicas do ambiente. Em contraste, os outros utilizam representações binárias, distinguindo apenas entre estados como “normal” e “assustado”, “comestível” e “ameaça” ou “caçando” e “caçado”, simplificando a percepção dos personagens quanto às condições de risco ou oportunidade representadas pelos fantasmas.

Outro elemento situacional que aparece está no A17, um estudo que trabalha com grafos e considera a existência de nós de junção (cruzamentos) no caminho entre o Pac-Man e os fantasmas.

No grupo de dados identificadores, inclui-se a distinção entre os fantasmas por tipo, representados por cor ou nome.

Finalmente, entre os dados de desempenho, encontram-se o total de passos percorridos pelo Pac-Man e a pontuação estimada por ação, utilizada em abordagens que preveem o valor esperado para o estado resultante de cada movimento, guiando a escolha da melhor alternativa.

Essas variáveis, embora heterogêneas em frequência e propósito, ilustram a diversidade de estratégias de codificação e representação dos dados. Em alguns casos, são utilizadas como entrada de redes neurais ou algoritmos de decisão. Em outros, constituem métricas para análises descritivas, comparativas ou preditivas, sendo indispensáveis para a compreensão e modelagem do comportamento no jogo Pac-Man.

3.2.2 Dados do Jogo

Os dados relacionados ao jogo foram identificados em 21 dos 25 estudos analisados e a Tabela 8 apresenta quais dados foram encontrados. A variável mais frequente é a pontuação (total, atual ou modificada), utilizada como métrica de desempenho, como sinal de recompensa ou como critério para decisões em aprendizado por reforço. Na

maioria dos casos, a pontuação corresponde diretamente às ações tradicionais do jogo, como comer pastilhas, derrotar fantasmas ou concluir níveis.

Tabela 8. Distribuição dos Dados do Jogo nos Artigos Selecionados

Dado do Jogo	Frequência	Artigos (Índice)
Pontuação (Total, Atual ou Modificada)	15	A01; A05; A06; A07; A08; A10; A11; A14; A16; A17; A18; A19; A20; A21; A24
Pastilhas, Pastilhas-Poder, Frutas (Presença, Posição)	15	A08; A09; A10; A11; A13; A14; A15; A16; A18; A19; A20; A21; A22; A23; A24
Tempo de Jogo	7	A04; A06; A10; A11; A17; A18; A22
Nível Atual	3	A07; A22; A23
Vidas (Ganhas, Perdas, Restantes)	2	A07; A18
Resultado da Partida (Vitória / Derrota)	2	A01; A04
Eventos Customizados (Combos)	1	A16
Tempo de Exibição (Frutas)	1	A07
Estado de Completude	1	A06
Média de Tentativas para Completar	1	A23
Não se aplica	4	A02; A03; A12; A25

Fonte: Elaborado pelos autores (2025)

Em alguns estudos, no entanto, a pontuação foi modificada em relação às regras originais. Em A16, A17 e A19 foram ajustados os valores atribuídos à coleta de pastilhas normais, pastilhas de poder e à captura de fantasmas. Além desses elementos comuns, o A16 também alterou a pontuação associada à vitória, à derrota e à ativação de combos por eliminação sequencial de fantasmas. Já o A19 ajustou ainda os valores relativos à coleta de frutas e à penalidade por morte.

A presença ou posição de pastilhas, pastilhas de poder e frutas também é utilizada. Esses elementos são usados tanto para informar decisões locais — como distância até o próximo item — quanto para construir mapas parciais ou globais da distribuição de recompensas. O tempo de execução é utilizado majoritariamente como métrica de avaliação de desempenho, representando o tempo necessário para completar episódios ou atingir recompensas.

Outros dados extraídos dos jogos incluem o nível atual alcançado, o número de vidas ganhas, perdidas ou restantes, e o resultado da partida, considerando vitória ou derrota. Também foram utilizados o tempo de exibição de frutas no labirinto, o estado de completude do cenário e a média de tentativas necessárias para completar o labirinto.

Assim como nas demais categorias, os dados do jogo são codificados de maneiras diversas. Seu uso é especialmente relevante em análises de desempenho, planejamento estratégico, balanceamento de dificuldade e construção de simulações.

3.2.3 Dados Externos ao Jogo

Embora o foco principal da revisão esteja nos dados internos ao jogo, 5 estudos também incorporaram dados externos, geralmente relacionados a aspectos fisiológicos, perceptuais ou subjetivos dos participantes.

Tabela 9. Distribuição dos Tipos de Dados Externos nos Artigos Seleccionados

Dado Externo (Detalhado)	Frequência	Artigos (Índice)
Dados Demográficos (idade, sexo, profissão, tempo dedicado a jogos)	4	A03; A07; A12; A22
Autoavaliação Subjetiva (valência, excitação, frustração, dificuldade percebida, avaliação de desempenho)	4	A02; A03; A12; A22
Dados Fisiológicos (expressões faciais, ângulo da cabeça, sinais respiratórios)	2	A02; A03
Comentários Livres / Textuais	1	A12
Análise Visual/Ocular	1	A23
Não se aplica	19	A01; A04; A05; A06; A08; A09; A10; A11; A13; A14; A15; A16; A17; A18; A19; A20; A21; A24; A25

Fonte: Elaborado pelos autores (2025)

Esses dados foram usados para enriquecer as análises, permitindo correlações entre o comportamento no jogo e estados emocionais, cognitivos ou sociais. A Tabela 9 contém a relação dos dados achados. Esses dados externos foram frequentemente associados a segmentos específicos do *gameplay* por *timestamps* ou blocos sincronizados, possibilitando análises que relacionam evento no jogo e estado do jogador. Apesar de sua menor frequência, essas abordagens destacam o potencial de integração entre métricas internas e externas na avaliação do comportamento em jogos.

Embora estes dados tenham relevância científica, eles não são o foco desta revisão. Por isso não foi feita uma análise detalhada de seus aspectos e de seus resultados.

4. Discussão dos Resultados

Aqui apresenta-se a análise das questões de pesquisa propostas para esta revisão sistemática. A questão principal investigada foi “Como o jogo Pac-Man tem sido analisado em pesquisas científicas?”. O Pac-Man tem sido analisado por meio de diferentes formas de coleta de dados, principalmente automática, mas também por captura de tela, questionários e sensores externos, com organização em três grupos: dados dos personagens, dados do jogo e dados externos. Os estudos revelam diversidade de abordagens, porém sem padronização na descrição dos dados e com uso ainda limitado de análises sistemáticas entre variáveis internas.

Quanto aos métodos de coleta e tratamento, predominam métodos automáticos de coleta durante a execução do jogo, complementados por capturas de tela, *loggers*, questionários e sensores fisiológicos. O tratamento dos dados, quando relatado, restringe-se geralmente à organização para análise posterior.

No que se refere aos tipos de dados os estudos exploram dados dos personagens (posições, distâncias, ações, estados), do jogo (pontuação, itens, tempo, níveis, vidas) e externos (demográficos, subjetivos, fisiológicos, visuais). Essa variedade amplia as possibilidades de análise, mas aparece com heterogeneidade de detalhamento entre os trabalhos.

Por fim, as técnicas de análise recorrem sobretudo à inteligência artificial: regras, busca em árvore, algoritmos evolucionários, aprendizado por reforço e redes neurais. O uso de testes estatísticos clássicos é pontual, em geral restrito a comparações ou validações, sendo rara a investigação sistemática de relações entre variáveis internas do jogo.

Além da análise das questões de pesquisa, registra-se que, durante a busca preliminar de estudos, foi encontrada uma revisão conduzida por Rohlfshagen et al. (2017), que apresenta um levantamento abrangente sobre o uso dos jogos Pac-Man e Ms. Pac-Man em pesquisas acadêmicas ao longo de duas décadas. Embora se trate de um estudo secundário e, portanto, não incluído diretamente no corpus da presente revisão sistemática, seus achados permitiram uma comparação útil com os resultados aqui obtidos. Embora tenham investigado estudos num período de 20 anos, o foco daquela revisão foi a área de inteligência computacional e portanto, mais específica.

Em síntese, a presente revisão sistematiza “quem coleta o quê e como” (tipos de agente e métodos de coleta), quais dados são efetivamente utilizados (personagens, jogo e externos) e quais técnicas são aplicadas aos dados (técnicas de IA/controle e uso pontual de análises estatísticas). Ao mesmo tempo, os resultados obtidos evidenciam a heterogeneidade de diversos aspectos de análise além de limitações na exploração estatística sistemática entre variáveis internas do Pac-Man.

5. Considerações Finais

A análise dos dados coletados nos estudos revisados revela uma série de variáveis amplamente utilizadas, sobretudo relacionadas aos estados dos personagens e a elementos básicos do ambiente. No entanto, considerando a estrutura e a dinâmica do jogo Pac-Man, é possível identificar outros dados potencialmente relevantes que não foram explicitamente considerados ou registrados nos artigos analisados.

Entre os estados e atributos de objetos do ambiente, destacam-se variáveis associadas às pastilhas, pastilhas de poder e frutas. Além da presença e posição, poderiam ser registrados o estado de consumida ou não (variável binária), a quantidade restante no labirinto, ou ainda o tempo desde o consumo, permitindo análises temporais de progresso. No caso das frutas, além do estado ativo ou inativo, o tempo de exibição ou tempo restante para desaparecimento constituem variáveis não muito exploradas, bem como o valor específico de pontuação associado a cada tipo de fruta, que enquanto pontualmente especificado, não aparece como informação utilizada.

Em relação às distâncias, poderiam ser consideradas métricas como a distância até a próxima pastilha, a distância até o centro do labirinto, a distância entre fantasmas (relevante para estratégias de dispersão ou colaboração) e a distância até cruzamentos, que podem influenciar decisões estratégicas e planejamento tático dos personagens.

No que se refere às informações temporais, o tempo restante de vulnerabilidade dos fantasmas após o consumo de uma pastilha de poder, bem como o tempo até o reaparecimento de frutas, que podem ser determinantes para agentes que priorizam maximizar a pontuação. Além disso, poderiam ser computados o tempo desde o último evento relevante (como o consumo de uma pastilha de poder ou a última morte do Pac-Man) e o tempo de exposição a risco, representando quanto tempo o Pac-Man permaneceu

próximo de um ou mais fantasmas em condições de ameaça. Também poderiam ser consideradas a densidade local de pastilhas (número de pastilhas próximas dentro de um raio específico), a frequência de trocas de direção por unidade de tempo (indicativa do estilo de navegação do personagem) e o padrão de uso de teletransportes.

Adicionalmente, variáveis derivadas do comportamento dos personagens podem enriquecer a modelagem, como a proximidade média aos fantasmas durante a partida, o número de confrontos bem-sucedidos (quantidade de fantasmas comidos após o consumo de uma pastilha de poder) e a efetividade no uso dessas pastilhas, expressa pela razão entre o número de fantasmas comidos e o número de pastilhas de poder consumidas. Outras métricas de desempenho poderiam incluir a razão entre pontos obtidos e tempo gasto (medida de eficiência), a média de distância percorrida entre eventos relevantes (como entre o consumo de uma pastilha de poder e o primeiro fantasma comido) e a quantidade de pontos desperdiçados, representada por frutas ou fantasmas vulneráveis que desapareceram sem serem coletados.

Por fim, apesar da diversidade de artigos encontrados, algumas limitações são observadas. Uma delas diz respeito ao nível de detalhamento na descrição dos dados utilizados. Em muitos casos, a identificação dos dados e métodos de tratamento depende da leitura de exemplos, trechos de código ou descrições indiretas, o que compromete a transparência e dificulta tanto a replicação quanto a comparação entre os estudos. Também se observa que, embora análises estatísticas clássicas como testes *t* ou correlações sejam utilizadas em alguns estudos, elas ainda são aplicadas de forma pontual e restrita a comparações entre modelos ou validações descritivas. Raramente são utilizadas para investigar relações entre variáveis específicas do jogo—como decisões, estados e resultados—de maneira sistemática. Esse é um ponto que pode ser explorado com mais profundidade, sobretudo quando combinadas a técnicas de visualização e exploração de padrões em dados de *gameplay*. Como trabalhos futuros, há a possibilidade de ampliar o período de análise para avaliar estudos anteriores ou ainda incluir outras bases de dados para verificar se novos estudos relevantes podem ser encontrados e analisados pelos mesmos critérios.

Agradecimentos

Os autores agradecem a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, à UDESC, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de produtividade DT2, processo 306613/2022-0, e à Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC) pelo financiamento parcial ao laboratório LARVA (Laboratory for Research on Visual Applications) e ao laboratório do GPIE (Grupo de Pesquisa em Informática na Educação) T.Os. No.: 2023TR284 e 2023TR000245.

Referências

- An, J. and Du, Y. (2022). Training agent to play pac-man under authentic environment based on image recognition. In *2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*, pages 68–75. IEEE.
- Banerjee, T., Khasnobish, A., Chowdhury, A., and Chatterjee, D. (2019). Reckoning res- piratory signals to affectively decipher mental state. In *2019 41st Annual*

- International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4654–4659. IEEE.
- Blom, P. M., Bakkes, S., and Spronck, P. (2019). Modeling and adjusting in-game difficulty based on facial expression analysis. *Entertainment Computing*, 31:100307.
- Buchinger, D., de Siqueira Cavalcanti, G. A., and da Silva Hounsell, M. (2014). Mecanismos de busca acadêmica: uma análise quantitativa. *Revista brasileira de computação aplicada*, 6(1):108–120.
- Butt, N. A., Mahmood, Z., Rehman, G. U., Nasralla, M. M., Zubair, M., Farman, H., and Khattak, S. B. A. (2023). The development of intelligent agents: A case-based reasoning approach to achieve human-like peculiarities via playback of human traces. *IEEE Access*, 11:78693–78712.
- Cai, J., Li, J., Zhang, M., and Tei, K. (2023). Value iteration networks with gated summarization module. *IEEE Access*, 11:60407–60420.
- Cowley, B. and Charles, D. (2016). Behavlets: a method for practical player modelling using psychology-based player traits and domain specific features. *User Modeling and User-Adapted Interaction*, 26:257–306.
- Costa, D. B., Lucca, G., and Adamatti, D. F. (2022). Aprendizado por reforço e jogos: uma proposta focada na análise de algoritmos. *Revista Brasileira de Computação Aplicada*, 14(2):26–34.
- Deterding, S. (2015). The lens of intrinsic skill atoms: A method for gameful design. *Human-Computer Interaction*, 30:294–335.
- Drachen, A., El-Nasr, M. S., and Canossa, A. (2013). *Game analytics: maximizing the value of player data*. Springer.
- Fachantidis, A., Taylor, M. E., and Vlahavas, I. (2017). Learning to teach reinforcement learning agents. *Machine Learning and Knowledge Extraction*, 1(1):21–42.
- Foderaro, G., Swinger, A., and Ferrari, S. (2016). A model-based approach to optimizing ms. pac-man game strategies in real time. *IEEE Transactions on Computational Intelligence and AI in Games*, 9(2):153–165.
- Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, 1(1):20–20.
- Green, C. S. and Bavelier, D. (2012). Learning, attentional control, and action video games. *Current biology*, 22(6):R197–R206.
- Hatschka, C., Ciabattoni, A., and Eiter, T. (2023). Deontic paradoxes in asp with weak constraints. *arXiv preprint arXiv:2308.15870*.
- Johnson, D., Gardner, J., and Wiles, J. (2004). Experience as a moderator of the media equation: the impact of flattery and praise. *International Journal of Human-Computer Studies*, 61(3):237–258.
- Kitchenham, B., Charters, S., et al. (2007). Guidelines for performing systematic literature reviews in software engineering.

- Maddipati, H., Kundurthi, A., Raaj, P., Srilatha, K., and Surapaneni, R. (2020). Artificial intelligence based pacman game. *International Journal of Innovative Technology and Exploring Engineering*, 9:140–144.
- Mahlmann, T., Drachen, A., Togelius, J., Canossa, A., and Yannakakis, G. N. (2010). Predicting player behavior in tomb raider: Underworld. In *Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games*, pages 178–185. IEEE.
- Miranda, M., Sa´nchez-Ruiz, A. A., and Peinado, F. (2021). Interactive explainable case- based reasoning for behavior modelling in videogames. In *2021 IEEE 33rd Inter- national Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1263–1270. IEEE.
- Morosan, M. and Poli, R. (2017). Evolving a designer-balanced neural network for ms pacman. In *2017 9th Computer Science and Electronic Engineering (CEECE)*, pages 100–105. IEEE.
- Neufeld, E. A., Bartocci, E., Ciabattini, A., and Governatori, G. (2022). Enforcing ethical goals over reinforcement-learning policies. *Ethics and Information Technology*, 24(4):43.
- Rohlfshagen, P., Liu, J., Perez-Liebana, D., and Lucas, S. M. (2017). Pac-man conquers academia: Two decades of research using a classic arcade game. *IEEE Transactions on Games*, 10(3):233–256.
- Safak, A. B., Bostanci, E., and Soylucicek, A. E. (2016). Automated maze generation for ms. pac-man using genetic algorithms. *International Journal of Machine Learning and Computing*, 6(4):226–240.
- Santos Junior, E. S., de Oliveira Machado, A., Macedo, M. C., and dos Santos Souza, A. C. (2019). Reinforcement learning para treino do pac-man em speedrun. *Brazilian Journal of Development*, 5(11):25927–25957.
- Schrum, J. and Miikkulainen, R. (2015). Discovering multimodal behavior in ms. pac-man through evolution of modular neural networks. *IEEE transactions on computational intelligence and AI in games*, 8(1):67–81.
- Silva, G. A. and Rigo Ju´nior, L. O. (2020). Ensino de agentes inteligentes por meio de problemas em jogos. *Research, Society and Development*, 9(1):e129911793–e129911793.
- Sledge, I. J. and Pr´ncipe, J. C. (2018). Analysis of agent expertise in ms. pac-man using value-of-information-based policies. *IEEE Transactions on Games*, 11(2):142–158.
- Tamassia, M., Zambetta, F., Raffe, W. L., Mueller, F., and Li, X. (2017). Learning options from demonstrations: A pac-man case study. *IEEE Transactions on Games*, 10(1):91–96.
- Wang, B., Ma, R., Kuang, J., and Zhang, Y. (2020). How decisions are made in brains: Unpack “black box” of cnn with ms. pac-man video game. *IEEE Access*, 8:142446–142458.
- Wolterink, M. and Bakkes, S. (2021). Towards explainable prediction of player frustration in video games. In *Proceedings of the 16th International Conference on the Foundations of Digital Games*, pages 1–10.

- Yang, Q., Lin, Z., Zhang, W., Li, J., Chen, X., Zhang, J., and Yang, T. (2022). Monkey plays pac-man with compositional strategies and hierarchical decision-making. *Elife*, 11:e74500.
- Yee, N. (2006). The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments. *Presence: Teleoperators and virtual environments*, 15(3):309–329.
- Zhang, Q., Yao, J., Yin, Q., and Zha, Y. (2018). Learning behavior trees for autonomous agents with hybrid constraints evolution. *Applied Sciences*, 8(7):1077.
- Zikky, M. (2016). Review of a*(a star) navigation mesh pathfinding as the alternative of artificial intelligent for ghosts agent on the pacman game. *EMITTER International journal of engineering technology*, 4(1):141–149.