

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA
E TECNOLOGIA DO TRIÂNGULO MINEIRO

Vinícius Silveira Bisinoto

**Reconhecimento Automático de Valores em Dados de
RPG Utilizando Redes Neurais Convolucionais**

UBERABA - MG
2025

Vinícius Silveira Bisinoto

Reconhecimento Automático de Valores em Dados de RPG Utilizando Redes Neurais Convolucionais

Trabalho de Conclusão de Curso apresentado ao Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro, Campus Uberaba Parque Tecnológico, como requisito parcial para a conclusão do Curso de Engenharia da Computação.

Orientador: Prof. Dr. Marcelo Barreiro

Sumário

1	Introdução	2
2	Fundamentação Teórica	3
2.1	Role-Playing Games (RPGs)	3
2.2	Dados em Jogos de RPG	3
2.3	Visão Computacional	4
2.4	Reconhecimento de Padrões	4
2.5	Aprendizado de Máquina e Redes Neurais Artificiais	4
2.6	Redes Neurais Convolucionais (CNNs)	5
2.7	Aplicações Relevantes de CNNs	5
3	Metodologia	6
3.1	Ambiente de Desenvolvimento	6
3.2	Base de Dados Inicial	6
3.3	Arquitetura da Rede Neural Convolucional	6
3.4	Planejamento da Base de Dados de RPG	7

1 Introdução

Os jogos de interpretação de papéis (Role-Playing Games – RPGs) têm ganhado espaço significativo não apenas como forma de entretenimento, mas também como ferramenta pedagógica, terapêutica e social. Uma característica marcante desses jogos é o uso de dados físicos para determinar o sucesso ou fracasso de ações, contribuindo para a imersão, a aleatoriedade genuína e o envolvimento emocional dos jogadores [1, 2].

Apesar do crescimento de plataformas digitais que oferecem alternativas de rolagem automatizada, muitos jogadores ainda preferem o uso dos dados físicos. Essa preferência se deve tanto ao valor simbólico e emocional atribuído aos objetos quanto à percepção de que a aleatoriedade física é mais justa ou autêntica. Esse vínculo subjetivo com os dados físicos reflete não apenas uma tradição cultural, mas também uma experiência sensorial significativa que contribui para a ambientação e imersão no jogo.

Contudo, essa escolha gera um desafio prático. Em sessões que envolvem múltiplas rolagens — como testes de dano ou ações conjuntas — a contagem manual dos resultados pode se tornar demorada e propensa a erros. Esse tempo de cálculo afeta diretamente o ritmo narrativo da partida, principalmente quando há muitos jogadores na mesa ou situações de combate simultâneo. Mesmo entre grupos experientes, a perda de fluidez entre a ação e a contagem numérica pode comprometer a imersão da experiência.

Neste contexto, este trabalho propõe o desenvolvimento de um sistema capaz de reconhecer automaticamente os valores das faces superiores de dados de RPG físicos, a partir de imagens capturadas em tempo real. A proposta envolve o uso de técnicas de visão computacional e aprendizado profundo, com foco na aplicação de Redes Neurais Convolucionais (CNNs), conhecidas por sua eficiência no reconhecimento de padrões visuais em imagens [3, 4, 1].

As CNNs têm se destacado em áreas como reconhecimento facial, diagnóstico por imagem, leitura de placas veiculares e OCR (Reconhecimento Óptico de Caracteres), sendo aplicadas com sucesso em contextos que exigem robustez, acurácia e capacidade de generalização [3, 5]. Sua estrutura em camadas permite a extração hierárquica de características visuais, o que torna essa arquitetura ideal para o reconhecimento de símbolos, números ou formas em objetos tridimensionais como os dados de RPG, mesmo sob condições variadas de iluminação, ângulo ou resolução.

A motivação principal para este estudo decorre da própria experiência do autor enquanto mestre de RPG, tendo observado a frequência com que os jogadores optam pelos dados físicos mesmo diante de soluções digitais. Essa escolha, embora compreensível, introduz uma fricção significativa no andamento das sessões, especialmente quando são rolados diversos dados simultaneamente. Automatizar a leitura visual desses dados representa uma solução que respeita a tradição do jogo físico e, ao mesmo tempo, introduz uma camada de eficiência tecnológica à experiência.

Essa proposta também dialoga com avanços recentes na integração de tecnologias assistivas e aplicações de visão computacional em tempo real. Sistemas de OCR, por exemplo, já são amplamente utilizados para converter textos em papel para formatos digitais editáveis. Da mesma forma, um sistema capaz de interpretar os resultados de dados físicos em tempo real pode se tornar uma ferramenta de apoio para mestres e jogadores, especialmente em partidas online ou híbridas, onde o compartilhamento visual dos dados é necessário.

Dessa forma, o objetivo geral deste trabalho é desenvolver e treinar um modelo baseado em redes neurais convolucionais capaz de reconhecer, com alto grau de acurácia, os valores

apresentados nas faces superiores de dados de RPG a partir de imagens. Para isso, propõe-se:

- Elaborar um conjunto de dados composto por imagens de dados de RPG físicos, simulando diferentes condições de iluminação, ângulo e tipo de dado;
- Projetar e treinar uma arquitetura de CNN apropriada à tarefa de classificação visual dos dados;
- Avaliar a acurácia e o desempenho do modelo em situações práticas simuladas;
- Validar o potencial de aplicação da tecnologia em mesas de RPG reais.

A proposta aqui desenvolvida parte, portanto, de uma necessidade prática observada em mesas de RPG e da possibilidade de utilizar técnicas modernas de aprendizado profundo para otimizar o reconhecimento visual de dados físicos. Ao longo do trabalho, são apresentados os fundamentos teóricos necessários, a metodologia de desenvolvimento adotada, os resultados obtidos e, por fim, as conclusões e perspectivas futuras que surgem a partir desta iniciativa.

2 Fundamentação Teórica

Este capítulo apresenta os principais conceitos e fundamentos que sustentam o desenvolvimento deste trabalho. Serão abordados os aspectos históricos e técnicos do RPG, os tipos de dados utilizados nesses jogos, os fundamentos de visão computacional, o reconhecimento de padrões, as redes neurais artificiais e, em especial, as redes neurais convolucionais (CNNs), tecnologia central desta pesquisa.

2.1 Role-Playing Games (RPGs)

Os jogos de interpretação de papéis, conhecidos como Role-Playing Games (RPGs), são uma forma de jogo narrativo em que os participantes assumem papéis fictícios e interagem em um mundo imaginário, muitas vezes guiados por um narrador ou mestre [2]. Esses jogos ganharam notoriedade não apenas como forma de lazer, mas também como ferramentas de ensino, inclusão social e desenvolvimento emocional, ao promoverem criatividade, resolução de problemas, empatia e trabalho em equipe.

Além disso, os RPGs físicos — aqueles que envolvem livros, fichas e dados — permanecem como uma das formas mais tradicionais e culturalmente ricas de interação nesse gênero, mesmo em um cenário atual fortemente digitalizado. A experiência tátil, a interpretação ao vivo e a imprevisibilidade dos dados físicos contribuem para uma experiência singular e imersiva, valorizada por seus praticantes.

2.2 Dados em Jogos de RPG

Diferentemente dos jogos de tabuleiro convencionais, os RPGs utilizam uma variedade de dados poliédricos para simular testes de atributos e ações. Entre os tipos mais comuns estão o d4 (tetraedro), d6 (cubo), d8 (octaedro), d10 (decágono), d12 (dodecágono) e d20 (icosaedro), sendo cada um empregado conforme a mecânica do sistema de jogo.

A leitura dos valores apresentados nas faces superiores desses dados é fundamental para o progresso da narrativa. No entanto, essa tarefa torna-se complexa em sessões com múltiplas rolagens simultâneas. Em especial, sistemas que utilizam muitos dados para calcular danos, acertos ou testes em grupo enfrentam perda de tempo significativa, além de suscetibilidade a erros na leitura manual.

Essa dificuldade cresce conforme a diversidade de dados utilizados. Dados com faces pequenas ou com símbolos estilizados, como os d4 e alguns d10, tornam a interpretação visual ainda mais desafiadora, o que reforça a necessidade de soluções que automatizem essa leitura sem comprometer a fidelidade da experiência física.

2.3 Visão Computacional

A visão computacional é uma área da inteligência artificial que busca desenvolver métodos para que sistemas computacionais possam interpretar e compreender informações visuais do mundo real [6]. Isso envolve desde a captação de imagens, passando por seu processamento e análise, até a extração de dados relevantes para tomada de decisão automatizada.

No presente trabalho, a visão computacional é empregada com o intuito de detectar, identificar e interpretar os valores apresentados na face superior dos dados físicos de RPG. Isso requer o domínio de técnicas como pré-processamento de imagens, filtragem de ruído, segmentação, detecção de contornos e extração de características visuais que serão posteriormente classificadas por modelos baseados em aprendizado profundo.

2.4 Reconhecimento de Padrões

O reconhecimento de padrões é o campo responsável por identificar regularidades e estruturas em dados brutos, com o objetivo de classificá-los ou agrupá-los de forma significativa [7]. Ele é amplamente aplicado em tarefas como reconhecimento facial, leitura de textos (OCR), análise de expressões faciais e identificação de objetos em imagens.

No contexto deste trabalho, o reconhecimento de padrões é a base para que se possa distinguir, de forma automática, os valores apresentados nas faces dos dados físicos. Essa tarefa exige que o sistema seja capaz de lidar com variações de iluminação, ângulo, foco e tipo de fonte ou símbolo presente nos dados, o que torna indispensável o uso de modelos robustos e adaptáveis, como as redes neurais convolucionais.

2.5 Aprendizado de Máquina e Redes Neurais Artificiais

O aprendizado de máquina (machine learning) é uma subárea da inteligência artificial focada no desenvolvimento de algoritmos capazes de aprender padrões e realizar previsões ou classificações com base em dados. Entre os modelos mais poderosos e amplamente utilizados nessa área estão as redes neurais artificiais [5].

Essas redes são compostas por camadas de unidades chamadas neurônios artificiais, que processam informações em conjunto. Por meio do treinamento supervisionado, elas ajustam seus parâmetros internos com base em conjuntos de dados rotulados, adquirindo a capacidade de realizar inferências em dados novos e não vistos.

Tais características fazem das redes neurais ferramentas valiosas em aplicações complexas, como reconhecimento de fala, tradução automática, diagnósticos médicos e, principalmente, análise de imagens.

2.6 Redes Neurais Convolucionais (CNNs)

As Redes Neurais Convolucionais (Convolutional Neural Networks — CNNs) são uma arquitetura especializada de redes neurais artificiais, voltada especificamente ao processamento de imagens e dados com estrutura espacial [3]. Sua principal característica é a capacidade de extrair automaticamente padrões relevantes em imagens, sem necessidade de pré-processamento manual.

As CNNs operam por meio de camadas convolucionais que aplicam filtros (ou kernels) sobre a imagem de entrada, capturando características locais como bordas, formas e texturas. Essas informações são progressivamente refinadas ao longo das camadas, permitindo que o modelo aprenda representações hierárquicas de alta complexidade.

Além das camadas convolucionais, as CNNs costumam incluir:

- **Camadas de Pooling:** responsáveis por reduzir a dimensionalidade das representações intermediárias, agregando informações locais e promovendo invariância a pequenas mudanças na entrada;
- **Camadas Densas:** totalmente conectadas, utilizadas para combinar as características extraídas e realizar a classificação final;
- **Camada de Saída:** fornece a predição do modelo, que pode assumir diferentes formatos conforme a tarefa (classificação binária, multiclasse, etc.).

O uso de CNNs neste trabalho permite que o sistema reconheça os dígitos nas faces superiores dos dados de RPG com alto grau de robustez, mesmo em condições adversas, como iluminação irregular, oclusão parcial ou variações na distância da câmera.

2.7 Aplicações Relevantes de CNNs

As redes neurais convolucionais têm sido empregadas com sucesso em uma ampla gama de aplicações modernas. Dentre as mais relevantes, destacam-se:

- **Reconhecimento Facial:** utilizado em autenticação biométrica e sistemas de vigilância;
- **Diagnóstico Médico por Imagem:** como na detecção de tumores em exames de imagem, por exemplo, em mamografias ou tomografias;
- **OCR (Reconhecimento Óptico de Caracteres):** leitura automática de documentos e placas de veículos;
- **Sistemas de Navegação e Robótica:** reconhecimento de sinais visuais em ambientes dinâmicos;
- **Indústria e Logística:** inspeção automatizada de produtos e leitura de códigos visuais.

Tais aplicações comprovam a versatilidade e eficácia das CNNs, e embasam sua escolha como tecnologia central para a proposta deste TCC, que busca integrar o potencial do aprendizado profundo ao contexto lúdico dos jogos de RPG.

3 Metodologia

Esta seção descreve os procedimentos adotados para o desenvolvimento do modelo de reconhecimento automático de valores em dados de RPG utilizando redes neurais convolucionais (CNNs). Inicialmente, foram conduzidos testes com uma base de dados pública, composta por imagens de cães e gatos, a fim de validar a arquitetura da rede e o fluxo de processamento. Em seguida, planejou-se a criação de uma base própria com dados de RPG, respeitando aspectos como visibilidade das faces e condições variadas de iluminação.

3.1 Ambiente de Desenvolvimento

O modelo foi desenvolvido utilizando a biblioteca TensorFlow, com a API `tf.keras`, em um ambiente baseado em Jupyter Notebook. O código foi escrito em Python 3. A execução dos experimentos foi realizada em uma máquina com suporte a GPU, utilizando bibliotecas auxiliares como NumPy, Matplotlib e scikit-learn.

3.2 Base de Dados Inicial

Para os testes preliminares, utilizou-se uma base pública contendo imagens rotuladas de cães e gatos, com distribuição balanceada entre as classes. Essa base serviu para a validação da arquitetura inicial da CNN e do pipeline de treinamento. As imagens foram redimensionadas para 180×180 pixels, com normalização dos valores de pixel no intervalo $[0, 1]$.

3.3 Arquitetura da Rede Neural Convolucional

A rede neural convolucional desenvolvida neste trabalho foi projetada para realizar tarefas de classificação de imagens com eficiência e baixa complexidade. A seguir, apresenta-se a descrição da arquitetura utilizada para o problema inicial de classificação binária (gato ou cachorro).

- **Pré-processamento:** A primeira camada realiza a normalização dos valores dos pixels, convertendo-os para o intervalo $[0, 1]$ através da operação `Rescaling(1./255)`.
- **Camada Convolucional 1:** Conv2D com 16 filtros, tamanho de kernel 3×3 , ativação ReLU e preenchimento `same`, mantendo a dimensão espacial.
- **Max Pooling 1:** MaxPooling2D, responsável por reduzir a dimensionalidade da imagem.
- **Camada Convolucional 2:** Conv2D com 32 filtros e mesmas configurações da primeira camada convolucional.
- **Max Pooling 2:** Segunda operação de MaxPooling2D.
- **Camada Convolucional 3:** Conv2D com 64 filtros.
- **Max Pooling 3:** Redução final da dimensionalidade espacial.
- **Camada de Achplanamento:** Flatten, transforma a saída tridimensional em um vetor unidimensional.

- **Camada Densa Oculta:** Dense com 128 unidades e ativação ReLU.
- **Camada de Saída:** Dense(1) com ativação `sigmoid`, utilizada para classificar a imagem entre duas classes.

A rede foi compilada com os seguintes parâmetros:

- **Otimizador:** Adam, com taxa de aprendizado de 0,001.
- **Função de perda:** `binary_crossentropy`.
- **Métrica:** Acurácia.

A Figura 1 apresenta um diagrama ilustrativo da arquitetura da CNN descrita.



Figura 1: Arquitetura da CNN para classificação binária de imagens.

3.4 Planejamento da Base de Dados de RPG

O próximo passo do projeto envolve a construção de uma base de imagens contendo fotos reais de dados de RPG, priorizando faces de fácil visualização. Decidiu-se evitar o uso de dados tetraédricos (como o d4), pois a face superior nem sempre é bem definida visualmente. Pretende-se focar em dados como d6, d8, d10, d12 e d20, capturando múltiplas fotos de cada valor possível, em diferentes condições de iluminação e ângulos. A arquitetura será ajustada para lidar com múltiplas classes conforme o dado utilizado (por exemplo, 12 classes para um d12).

Referências

- [1] DATACAMP. Introdução às redes neurais convolucionais (cnns). <https://www.datacamp.com/pt/tutorial/introduction-to-convolutional-neural-networks-cnns>, 2023. Acesso em: 25 maio 2025.
- [2] HITCHENS, M.; DRACHEN, A. The many faces of role-playing games. *International Journal of Role-Playing*, v. 1, p. 3–21, 2007.
- [3] LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2324, 1998.
- [4] IBM. O que são redes neurais convolucionais? <https://www.ibm.com/br-pt/think/topics/convolutional-neural-networks>, 2024. Acesso em: 25 maio 2025.
- [5] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. Cambridge, MA: MIT Press, 2016. Tradução recomendada: Aprendizado profundo. Editora Elsevier, 2018.
- [6] SZELISKI, R. *Computer vision: Algorithms and applications*. London: Springer, 2010.
- [7] BISHOP, C. M. *Pattern recognition and machine learning*. New York: Springer, 2006.