

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

## Lista de Figuras

1	Exemplos de dados poliédricos utilizados em jogos de RPG. . . . .	4
2	Esquema conceitual da arquitetura de uma CNN, inspirado na LeNet-5 [1].	10
3	Trecho de código-fonte com a definição da arquitetura da CNN. . . . .	11

# Sumário

<b>1</b>	<b>Introdução</b>	<b>3</b>
<b>2</b>	<b>Fundamentação Teórica</b>	<b>6</b>
2.1	Role-Playing Games (RPGs) . . . . .	6
2.2	Dados em Jogos de RPG . . . . .	6
2.3	Visão Computacional . . . . .	7
2.4	Reconhecimento de Padrões . . . . .	7
2.5	Aprendizado de Máquina e Redes Neurais Artificiais . . . . .	7
2.6	Redes Neurais Convolucionais (CNNs) . . . . .	7
2.7	Aplicações Relevantes de CNNs . . . . .	8
<b>3</b>	<b>Metodologia</b>	<b>8</b>
3.1	Coleta e Estruturação do Conjunto de Dados . . . . .	9
3.2	Processamento de Imagens . . . . .	9
3.3	Divisão em Treinamento e Validação . . . . .	9
3.4	Arquitetura da Rede Neural Convolucional . . . . .	10
3.5	Treinamento do Modelo . . . . .	11

# 1 Introdução

Os jogos de interpretação de papéis (Role-Playing Games - RPGs) têm conquistado cada vez mais espaço não só como forma de entretenimento, mas também como ferramenta pedagógica, terapêutica e social. Desde o lançamento de *Dungeons & Dragons* em 1974, o RPG se consolidou como um gênero de jogos que une imaginação, cooperação e raciocínio lógico, promovendo experiências imersivas onde os jogadores atuam em conjunto na construção de narrativas. [2, 3]

Uma das características marcantes desses jogos é o uso de dados físicos para determinar sucessos ou fracassos de determinadas ações, como ataques, testes de habilidade ou eventos aleatórios, e, diferentemente dos jogos de tabuleiro convencionais, o RPG utiliza-se de uma diversidade de dados poliédricos — como d4, d6, d8, d10, d12 e d20 — que ampliam a variação probabilística das ações dos jogadores e tornam o jogo mais dinâmico e imprevisível. Além de sua usabilidade, esses dados se tornaram símbolos da cultura RPGista, muitas vezes personalizados com diferentes materiais, cores e gravuras, ou até eternizados em tatuagens, obras artísticas e coleções, reforçando o vínculo afetivo entre jogador e dado.[3]

Apesar do crescimento de plataformas digitais que oferecem alternativas de rolagem automatizada, como *Roll20* ou *Foundry VTT*, e até comandos de rolagem em bots de Discord, muitos jogadores ainda preferem o uso dos dados físicos. Essa preferência se deve tanto ao valor simbólico e afetivo atribuído aos objetos quanto à percepção de que a aleatoriedade física é mais justa ou autêntica do que a gerada computacionalmente. O som do dado rolando, o toque, o peso e até o suspense visual da face superior prestes a ser revelada são elementos sensoriais que contribuem para a ambientação do jogo e tornam a experiência mais sensorial, orgânica e memorável. [4]

No entanto, essa escolha estética e emocional traz um desafio prático. Em sessões que envolvem múltiplas rolagens com dados, que muitas vezes podem conter múltiplos dados simultâneos — como testes de dano, ataques em área ou ações conjuntas — a contagem manual dos resultados obtidos pode se tornar demorada e propensa a erros. Esse tempo de cálculo afeta diretamente o ritmo narrativo da partida, especialmente em jogos com muitos participantes ou quando os personagens utilizados estejam com suas habilidades já em maior nível, aumentando os números e resultando que dezenas de dados acabem sendo lançados ao mesmo tempo. Mesmo entre grupos experientes, a perda de fluidez entre ação e contagem numérica pode comprometer a imersão da experiência e gerar frustrações.

A crescente demanda por ferramentas que otimizem o tempo de jogo sem abrir mão da experiência de rolagem física motivou este trabalho, que propõe uma solução baseada em visão computacional para automatizar a leitura visual dos dados físicos durante uma rolagem, mantendo o aspecto tradicional do RPG com os dados físicos enquanto melhora a fluidez narrativa das sessões através da velocidade que um cálculo automático pode trazer. [5, 6, 7]

A Figura 1 ilustra alguns exemplos de dados utilizados em sistemas de RPG, destacando a variedade de formatos e faces numéricas. Essa diversidade, embora enriqueça o jogo, representa também uma dificuldade computacional para quem deseja automatizar a leitura desses elementos visuais.



Figura 1: Exemplos de dados poliédricos utilizados em jogos de RPG.

*Foto por Danielly Vitória Rodrigues, 2025.*

Partindo deste contexto, este trabalho tem como objetivo analisar a viabilidade do uso de Redes Neurais Convolucionais (CNNs) para o reconhecimento automático dos valores presentes nas faces superiores de dados físicos utilizados em jogos de RPG, a partir de um banco de imagens previamente separado e categorizado conforme o valor exibido na face superior dos dados. A proposta envolve o desenvolvimento e a avaliação experimental de um modelo de classificação multiclasse baseado em técnicas de visão computacional e aprendizado profundo, com foco na análise de seu desempenho em condições práticas simuladas. Para isso, será utilizada uma arquitetura de Rede Neural Convolucional (CNN) composta por múltiplas camadas convolucionais e de pooling, seguidas por camadas densas e uma camada de saída com função de ativação *softmax*, adequada ao reconhecimento de múltiplas classes de saída — correspondentes aos valores possíveis nas faces dos dados de RPG. As CNNs foram escolhidas por demonstrarem bastante eficiência em tarefas de reconhecimento de padrões visuais, especialmente em imagens, especialmente onde existem cenários com variação de iluminação, ângulo, cor e textura. [1, 8]

As Redes Neurais Convolucionais (CNNs) têm seu destaque em diversas áreas da inteligência artificial, como reconhecimento facial, diagnósticos por imagem, leitura de placas de veículos e OCR (Reconhecimento Óptico de Caracteres), sendo utilizadas amplamente devido ao seu sucesso em contextos que exigem robustez, acurácia e capacidade de generalização [1, 8]. Por ter várias camadas organizadas em etapas, essa arquitetura consegue

identificar detalhes visuais em diferentes níveis, o que a torna muito eficaz para reconhecer símbolos, números e formas em objetos tridimensionais. — como os dados físicos utilizados em jogos de RPG — mesmo sob condições variáveis de iluminação, ângulo ou resolução.

A principal motivação para o desenvolvimento deste estudo nasce da vivência pessoal do autor como mestre de RPG, especialmente ao observar, ao longo de diversas sessões, a forte preferência dos jogadores pelo uso de dados físicos. Mesmo com a ampla oferta de ferramentas digitais que simulam rolagens de forma prática e eficiente, é comum notar que muitos jogadores ainda optam pelos dados reais, seja pelo apego emocional, a sensação de adrenalina e expectativa ao rolar um dado, ou pela conexão simbólica que esses objetos representam dentro da cultura do RPG. No entanto, essa escolha, apesar de compreensível e até esperada, pode acabar trazendo desafios durante o jogo, principalmente em momentos em que múltiplos dados são rolados ao mesmo tempo. Nessas situações, a contagem manual dos resultados tende a consumir um tempo precioso, além de estar sujeita a erros que podem comprometer a precisão e, por consequência, a imersão e o ritmo narrativo da partida. Dessa forma, automatizar esse processo de leitura dos valores nos dados físicos surge como uma solução equilibrada: respeita e preserva a tradição dos jogos de mesa, mas aproveita os recursos da tecnologia para tornar a experiência mais fluida, dinâmica e eficiente para todos os participantes.

Essa proposta também se alinha com os avanços mais recentes no uso de sistemas de visão computacional, com destaque para o aumento da sua adesão em tecnologias assistivas. Um bom exemplo disso são os sistemas de reconhecimento óptico de caracteres (OCR), que hoje já são amplamente utilizados na conversão de documentos físicos em versões digitais editáveis, com uma precisão invejável. De forma semelhante, desenvolver um sistema capaz de identificar automaticamente os números presentes nas faces superiores dos dados físicos representa uma aplicação com potencial promissor. Tal recurso pode funcionar como uma importante ferramenta de apoio tanto para mestres quanto para jogadores, especialmente para aquelas sessões realizadas de forma online ou híbrida, onde a visualização e validação das rolagens físicas continua sendo valorizada como parte da experiência principal.

Com base nas necessidades identificadas ao longo da experiência prática com jogos de RPG de mesa, este trabalho tem como objetivo principal o desenvolvimento e a avaliação de um modelo computacional, fundamentado em redes neurais convolucionais (CNNs), capaz de reconhecer os valores exibidos nas faces superiores de dados físicos. Essa identificação será realizada a partir da análise de um conjunto de imagens previamente coletado e devidamente categorizado.

Para atingir esse objetivo, propõe-se então a execução das seguintes etapas:

- Construção de um banco de dados composto por imagens de diversos tipos de dados de RPG, capturadas sob diferentes ângulos de visão, exibindo valores diversos em sua face superior com o intuito de refletir a variabilidade presente em situações reais;
- Projeção e treinamento de uma arquitetura de rede neural convolucional voltada especificamente para a tarefa de classificação multiclasse dos valores numéricos visíveis nas imagens;
- Avaliação do desempenho do modelo utilizando métricas quantitativas, como acurácia e taxa de erro, considerando diferentes cenários simulados;

- Discussão voltada tanto à avaliação da aplicabilidade prática do sistema em sessões de RPG, quanto à identificação de pontos de melhoria e caminhos possíveis para o aprimoramento e a evolução do projeto em versões futuras.

A proposta citada parte, portanto, da oportunidade avaliada pela observação de uma limitação no uso dos dados físicos em sessões de RPG, propondo o uso de técnicas de aprendizado profundo como uma solução inovadora. Ao longo do trabalho, serão apresentados os fundamentos teóricos que embasam o projeto, a metodologia adotada para construção e validação do modelo, os resultados experimentais obtidos e, por fim, as considerações finais e sugestões para futuras melhorias.

DAQUI PRA BAIXO NÃO TEM NADA NOVO

## 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 [3]. 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 [5]. 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 [9]. 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 [8].

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 [1]. 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

Este capítulo descreve os materiais e os métodos empregados no desenvolvimento do presente trabalho, com foco na estruturação do conjunto de dados, no processamento das imagens, no modelo de rede neural convolucional utilizado e nas estratégias de treinamento e validação do classificador. As decisões adotadas ao longo do processo são justificadas com base na realidade do projeto e nas limitações observadas durante a execução experimental.

### 3.1 Coleta e Estruturação do Conjunto de Dados

Para a realização dos testes, foi construído um conjunto de imagens próprio contendo fotografias de dados físicos utilizados em jogos de RPG. As imagens foram capturadas em estúdio fotográfico, com auxílio da profissional Danielly Vitória Rodrigues, garantindo qualidade adequada de resolução, iluminação e enquadramento.

As imagens foram organizadas em pastas segundo os valores visíveis nas faces superiores dos dados, totalizando inicialmente 10 classes: *0\_zero* até *9\_nine*. Cada uma dessas pastas continha subpastas correspondentes aos diferentes tipos de dados (ex.: d4, d6, d8, d10). Essa estrutura permitiu manter a diversidade de formatos dentro de cada valor, simulando condições reais de uso.

Entretanto, durante a análise da distribuição dos dados, foi observado que as classes *0\_zero* até *9\_nine* possuíam um número significativamente inferior de imagens em relação às demais. Essa escassez comprometia a representatividade estatística e a robustez do treinamento, além de aumentar o risco de viés ou sobreajuste da rede.

Por essa razão, optou-se por restringir o escopo do modelo inicial às classes **1 a 8**, que apresentavam uma quantidade razoável de imagens (acima de 80 amostras por classe no total). As classes excluídas foram mantidas como possibilidade para expansão futura do projeto, mediante a ampliação da base.

### 3.2 Processamento de Imagens

Todas as imagens passaram por uma etapa de pré-processamento com o objetivo de remover espaços em branco ao redor dos dados e padronizar o formato das entradas. O processamento incluiu:

- **Corte absoluto:** Remoção de bordas fixas nas extremidades superior, inferior, esquerda e direita da imagem, mantendo foco no dado centralizado;
- **Redimensionamento:** Ajuste da imagem para dimensões fixas de 160x160 pixels, facilitando a entrada na rede convolucional;
- **Conversão RGB:** Padronização do canal de cor para compatibilidade com bibliotecas de visão computacional;

Esse tratamento foi realizado utilizando a biblioteca OpenCV. As imagens processadas foram salvas em pastas separadas para treino e validação.

### 3.3 Divisão em Treinamento e Validação

Para cada uma das classes (1 a 8), as imagens foram embaralhadas e divididas em 80% para treinamento e 20% para validação. A divisão foi feita com semente fixa para garantir reprodutibilidade. Devido ao tamanho limitado do dataset, optou-se por não subdividir o conjunto de validação em um conjunto adicional de teste. Essa decisão visa manter uma quantidade mínima de imagens por classe nas avaliações, evitando séries com menos de cinco amostras.

Para futuros trabalhos, recomenda-se a coleta de um novo conjunto independente de imagens para testes finais e validação cruzada do modelo.

### 3.4 Arquitetura da Rede Neural Convolucional

A arquitetura da Rede Neural Convolucional (CNN) utilizada neste trabalho foi projetada para equilibrar capacidade de extração de características e simplicidade estrutural, considerando a dimensão relativamente pequena do conjunto de dados e o objetivo de classificação multiclasse (valores de 1 a 8). A estrutura final consiste em:

- Uma camada de reescalonamento (Rescaling) que normaliza os valores dos pixels das imagens de entrada para o intervalo  $[0,1]$ , dividindo por 255. Essa etapa é essencial para acelerar a convergência durante o treinamento.
- Três blocos compostos por:
  - Uma camada Conv2D com filtros 3x3, ativação **ReLU** e preenchimento 'same', com 16, 32 e 64 filtros, respectivamente;
  - Uma camada *MaxPooling2D* após cada convolução, reduzindo as dimensões espaciais e promovendo generalização.
- Uma camada de *Flatten*, que transforma o mapa de características extraído em um vetor unidimensional.
- Uma camada densa *Dense* com 128 neurônios e ativação **ReLU**, responsável por combinar as características extraídas.
- Uma camada de saída *Dense* com 8 neurônios e ativação **softmax**, adequada para classificação multiclasse com rótulos inteiros de 0 a 7 (associados às classes 1 a 8).

O modelo foi compilado com otimizador Adam e função de perda *Sparse Categorical Crossentropy*, uma vez que os rótulos são inteiros e não codificados em one-hot, utilizando o otimizador Adam com taxa de aprendizado **0.0005**. A métrica de avaliação escolhida foi a acurácia.

A Figura 2 apresenta um diagrama ilustrativo inspirado na arquitetura LeNet-5, um dos modelos de redes neurais convolucionais mais clássicos, proposto por LeCun et al. (1998) para reconhecimento de dígitos manuscritos. Embora a estrutura exata da rede implementada neste trabalho difira em termos de parâmetros, profundidade e dimensionalidade das entradas, o diagrama é representativo do fluxo de processamento típico de uma CNN: camadas de convolução seguidas por subamostragem (pooling), camadas densas totalmente conectadas e uma camada final com função de ativação *softmax* para classificação multiclasse.

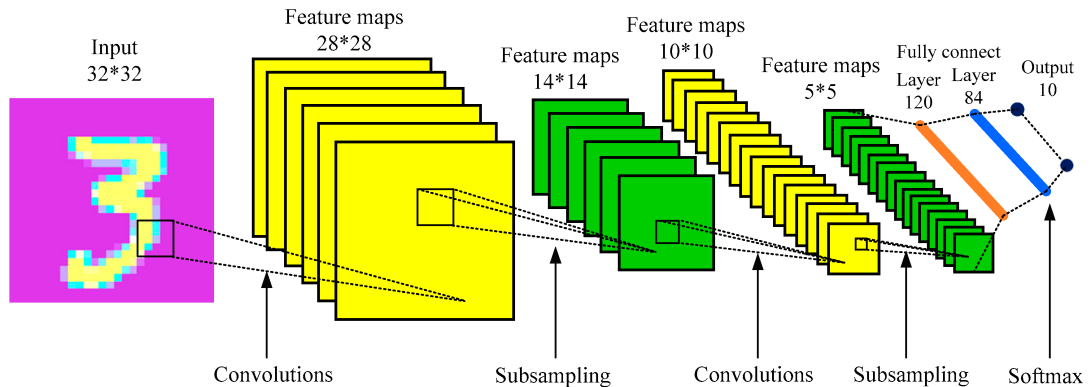


Figura 2: Esquema conceitual da arquitetura de uma CNN, inspirado na LeNet-5 [1].

A Figura 3 apresenta uma captura de tela do trecho de código-fonte responsável pela definição da arquitetura final da rede neural convolucional utilizada neste trabalho. A imagem evidencia a sequência de camadas adotadas no modelo, implementadas com a biblioteca Keras.

```
model = tf.keras.Sequential([
    tf.keras.layers.Rescaling(
        1. / image_color_channel_size,
        input_shape=image_shape
    ),
    tf.keras.layers.Conv2D(16, 3, padding='same', activation='relu'),
    tf.keras.layers.MaxPooling2D(),
    tf.keras.layers.Conv2D(32, 3, padding='same', activation='relu'),
    tf.keras.layers.MaxPooling2D(),
    tf.keras.layers.Conv2D(64, 3, padding='same', activation='relu'),
    tf.keras.layers.MaxPooling2D(),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(128, activation='relu'),
    tf.keras.layers.Dense(8, activation='softmax')
])
```

Figura 3: Trecho de código-fonte com a definição da arquitetura da CNN.

### 3.5 Treinamento do Modelo

O treinamento da rede neural convolucional (CNN) foi conduzido de forma iterativa, passando por diversas fases de ajuste de parâmetros, avaliação de desempenho e análise qualitativa dos resultados. O objetivo principal era obter um modelo capaz de classificar corretamente os valores nas faces superiores de dados físicos de RPG com alta acurácia e boa generalização.

Inicialmente, o modelo foi treinado por 20 épocas, com um *learning rate* de 0.001, sem aplicação de técnicas de *data augmentation*. Essa primeira configuração foi escolhida por simplicidade e por estar alinhada com abordagens comuns na literatura introdutória de CNNs. Apesar da acurácia de treino apresentar crescimento constante, a acurácia de validação mostrava valores baixos, sugerindo dificuldade de generalização do modelo.

Tentativas subsequentes foram feitas para aumentar a robustez do treinamento por meio de técnicas de aumento de dados. Utilizou-se uma camada de *data augmentation* composta por rotações, espelhamentos e zoom aleatórios. No entanto, todos os testes com essa abordagem resultaram em queda expressiva de desempenho, com redução da acurácia de validação para menos da metade do valor anteriormente obtido. Tais alterações possivelmente introduziram variações que descaracterizaram os padrões dos números nas faces dos dados, prejudicando o aprendizado.

Também foram avaliadas estratégias de tratamento de imagem, como conversão para escala de cinza, equalização de histograma e aplicação de *thresholding*. Embora visualmente eficazes em destacar o contraste entre o fundo branco e o dado, esses tratamentos acabaram por remover detalhes importantes dos números em certos casos. A perda de informação visual comprometeu a capacidade de discriminação do modelo, levando à decisão de manter as imagens em sua versão RGB original, apenas cortadas e redimensionadas para  $160 \times 160$  pixels.

Para melhorar a capacidade de aprendizado, o número de épocas foi aumentado para 40 e, posteriormente, 50. Com isso, a acurácia de treino superou os 90%, mas a acurácia de validação permaneceu significativamente inferior. Essa discrepância levantou a hipótese de um desequilíbrio na divisão dos conjuntos. O conjunto de validação, com apenas 10% das amostras (proporção 90/10), continha um número muito limitado de imagens por classe, o que impactava negativamente a avaliação do modelo.

A divisão foi então ajustada para 80/20 (treinamento/validação), o que resultou em melhoria expressiva. Em uma das execuções com 50 épocas, observou-se que a acurácia de validação estabilizava por volta da 40<sup>a</sup> época, formando um platô. Por isso, optou-se por reduzir o número de épocas para 40.

Além disso, o *learning rate* foi ajustado para 0.0005, com o intuito de proporcionar um aprendizado mais estável e menos sujeito a oscilações abruptas nos pesos da rede. Essa configuração final resultou em uma acurácia de treinamento de aproximadamente 84% e acurácia de validação de 72%, com perda de validação (*val\_loss*) estabilizada em torno de 0.95, o que indica um bom nível de generalização para a tarefa de classificação de oito classes distintas (valores de 1 a 8).

## Referências

- [1] 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.
- [2] PETERSON, J. *Playing at the world: A history of simulating wars, people and fantastic adventures, from chess to role-playing games*. San Diego, CA: Unreason Press, 2012.
- [3] HITCHENS, M.; DRACHEN, A. The many faces of role-playing games. *International Journal of Role-Playing*, v. 1, p. 3–21, 2007.
- [4] ZAGAL, J. P.; DETERDING, S. *Role-playing game studies: Transmedia foundations*. Routledge, 2017.
- [5] SZELISKI, R. *Computer vision: Algorithms and applications*. London: Springer, 2010.
- [6] IBM. O que são redes neurais convolucionais? <https://www.ibm.com/br-pt/think/topics/convolutional-neural-networks>, 2024. Acesso em: 25 maio 2025.
- [7] 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.
- [8] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. Cambridge, MA: MIT Press, 2016. Tradução recomendada: Aprendizado profundo. Editora Elsevier, 2018.
- [9] BISHOP, C. M. *Pattern recognition and machine learning*. New York: Springer, 2006.