

CLASSIFICAÇÃO DE ESPÉCIES FLORESTAIS USANDO APRENDIZAGEM EM PROFUNDIDADE (DEEP LEARNING)

FOREST SPECIES CLASSIFICATION BASED ON DEEP LEARNING

CLASIFICACIÓN DE ESPECIES FORESTALES USANDO APRENDIZAJE EN PROFUNDIDAD (DEEP LEARNING)

Skarlat Manuely da Silva de Souza¹
Marcelo Nepomoceno Kapp²

Resumo: O principal objetivo deste projeto é o desenvolvimento de um sistema para o reconhecimento automático de espécies florestais usando aprendizagem profunda (*Deep Learning*). Neste tipo de sistema, uma imagem digitalizada de uma madeira é utilizada como entrada e a espécie florestal da tábua é automaticamente informada como saída. Este trabalho está organizado como segue. Primeiramente, apresentamos as técnicas envolvendo a rede neural convolutiva utilizada. Em seguida, descrevemos os experimentos e análises realizadas comparando os atuais resultados com técnicas já estudadas anteriormente de extração de características e classificação. As avaliações serão conduzidas utilizando uma base de imagens de 112 espécies florestais distintas.

Palavras-chave: Aprendizagem profunda. Reconhecimento de espécies florestais. Redes Neurais Convolutivas.

Abstract: The main goal of this project is to develop a system for the automatic recognition of forest species using deep learning. In these systems, a scanned image of a wood is used as input and the wood species of the board is automatically informed as output. This paper is organized as follows. First, we present the techniques involving the convolutional neural network used. Next, we describe the experiments and compare the current results with previously studied techniques of extraction of characteristics and classification. The experiments will be conducted using a database of images of 112 different forest species.

Keywords: Deep learning. Recognition of forest species. Convolution neural networks.

Resumen: El principal objetivo de este proyecto es el desarrollo de un sistema para el reconocimiento automático de especies forestales usando aprendizaje profundo (*Deep Learning*). En este tipo de sistema, una imagen digitalizada de una madera se utiliza como entrada y la especie forestal de la tabla es automáticamente informada como salida. Este trabajo está organizado de la siguiente manera. Primero, presentamos las técnicas envolviendo la red neural convolutiva utilizada. A continuación, describimos los experimentos y análisis realizados comparando los actuales resultados con técnicas ya estudiadas anteriormente de extracción de características y clasificación. Las evaluaciones se llevarán a cabo utilizando una base de imágenes de 112 especies forestales distintas.

Palabras-clave: Aprendizaje profundo. Reconocimiento de especies forestales. Redes Neurales Convolutivas.

Envio: 25/02/2019

Revisão: 25/02/2019

Aceite: 27/05/2019

INTRODUÇÃO

A indústria madeireira lida diariamente com diferentes tipos de madeiras, extraídas de diferentes espécies florestais. A falta de fiscalização apropriada, entretanto, pode fazer com que madeiras de uma determinada espécie sejam comercializadas como se pertencessem a outra espécie. Por exemplo, espécies mais valorizadas podem ser trocadas por espécies de menor valor, o que pode causar prejuízo a um determinado comprador. Mais gravemente, o comércio de madeiras extraídas de espécies cuja comercialização é ilegal pode acarretar em um prejuízo ambiental ao país e ao planeta como um todo. A espécie de uma árvore pode ser, de certo modo, facilmente reconhecida quando a mesma encontra-se em seu habitat natural. Mas depois que a árvore é extraída e cortada em tábuas, muitas vezes apenas um especialista é capaz de afirmar de qual espécie a tábua foi extraída. Embora as autoridades governamentais estejam se esforçando cada vez mais para que haja um maior controle neste tipo de indústria, o controle apropriado da comercialização de madeiras dependeria da contratação e recrutamento de um grande número de especialistas. Como o orçamento dos governos é geralmente limitado, o reconhecimento automático de espécies florestais, realizado por computadores (como ilustrado na Figura 1), consiste em uma alternativa de menor custo para tratar este problema. O desenvolvimento deste tipo de sistema pode ser realizado dados avanços recentes relacionados a Visão Computacional [Forsyth & Ponce, 2002] e o Reconhecimento de Padrões (RP) [Duda & Hart, 2000].

Contudo, estas tarefas geralmente requerem um grande número de especialistas, o que demanda tempo e investimentos. Nesse contexto, o desenvolvimento de um sistema automatizado poderia reduzir os custos e tornar este processo mais rápido e eficaz. Vários sistemas foram propostos na literatura para o reconhecimento de espécies florestais por meio do desenvolvimento de softwares que sejam capazes de reconhecer uma espécie a partir da imagem capturada de uma madeira, auxiliando o ser humano. Neste estudo, modelos de aprendizagem profunda (*Deep Learning*) são investigados para melhorar a precisão em diferentes padrões de reconhecimento.

METODOLOGIA

No tipo de sistema que estudamos no projeto, uma imagem digitalizada de uma madeira é utilizada como entrada e a espécie florestal da tábua é automaticamente informada como saída (Figura 1). Uma das principais dificuldades é a implementação de um método eficiente de extração de características de texturas para tais imagens. Considerando isto, neste trabalho, são realizadas pesquisas que visam apresentar uma investigação de técnicas de aprendizagem profunda.

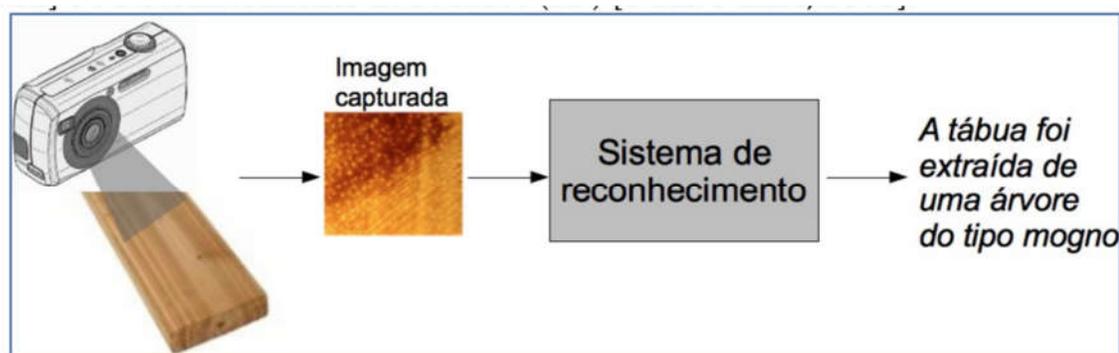


Figura 1: Concepção geral do sistema

Nessas técnicas utilizam-se dados brutos (por exemplo, pixels de imagem) como entrada para os modelos e deixa-se aprenderem representações intermediárias para domínios em que as principais características de formalizar para o reconhecimento. Suas abordagens inovadoras dispensam um pré-processamento e geram automaticamente propriedades invariantes da camada de entrada. Também possuem uma forte capacidade de realizar a classificação usando apenas um fragmento relativamente pequeno de texturas, o que é ideal para o problema de classificação abordado neste estudo. Nesse contexto, será utilizada uma arquitetura de rede neural convolutiva como técnica de Deep Learning. Abaixo descrevemos o método proposto em detalhes e todo o procedimento de implementação, já que rede neural convolutiva é composta por diferentes camadas, tais como:

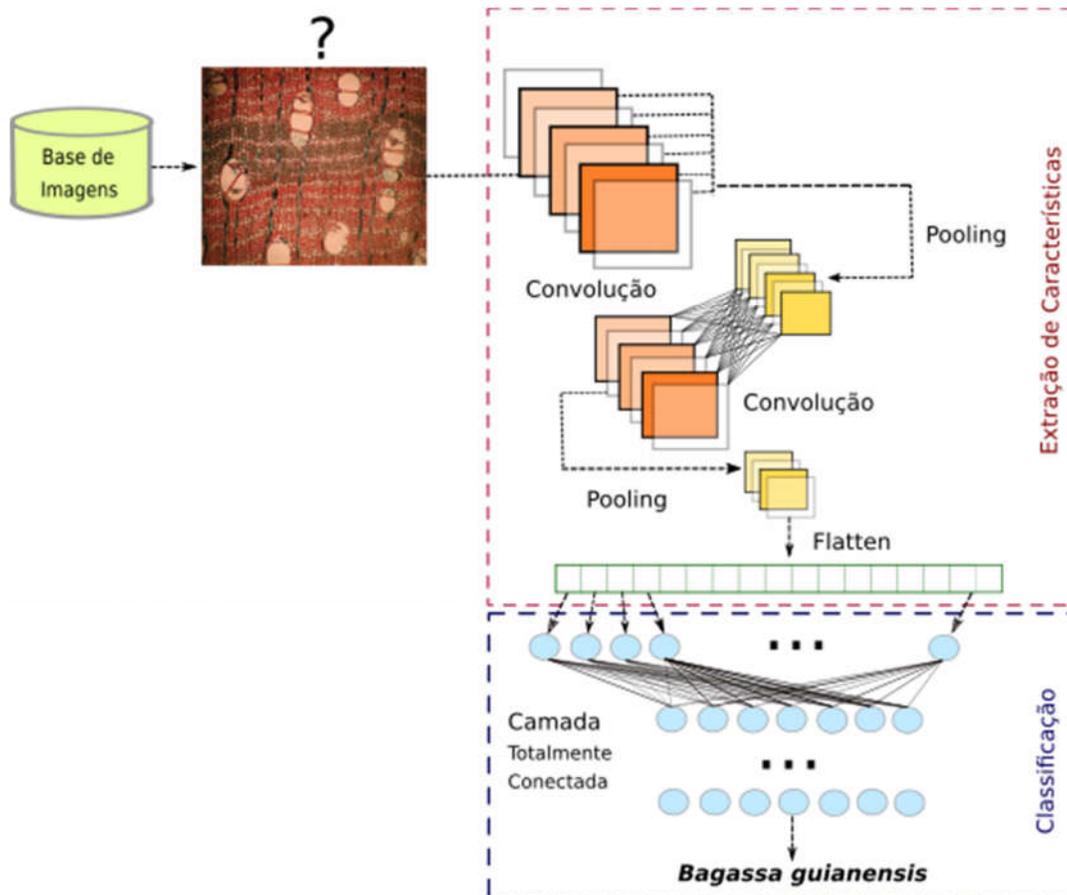
1) Camadas Convolucionais: Esta camada é composta por um conjunto de filtros (ou *kernels*) treináveis que são aplicados em toda a imagem. Esses filtros são convoluídos (através do

produto escalar) com os dados de entradas para obter um mapa de características. O filtro é percorrido por toda a imagem, sendo controlado pelo parâmetro conhecido como *stride*. Sobre o resultado da convolução é aplicada a função de ativação - ReLU (unidade de retificação linear), em cada elemento do resultado da convolução.

2) Camadas de Pooling: É utilizada com objetivo de reduzir o tamanho espacial das matrizes resultantes da convolução e capturar pequenas invariâncias com informações mais representativas. Consequentemente, essa técnica reduz a quantidade de parâmetros a serem aprendidos na rede, contribuindo para o controle de sobre ajuste (*overfitting*).

3) Camada totalmente conectada: É completamente conectada com a camada anterior. Como a camada totalmente conectada vêm após uma camada convolucional ou de pooling, é necessário conectar cada elemento das matrizes de saída de convolução em um neurônio de entrada. Contudo, para entrarem, as matrizes precisam passar por uma alteração em sua estrutura chamado *flatten* sendo colocados de forma linear para camada totalmente conectada. É nesta camada que o algoritmo de treinamento supervisionado *backpropagation* é empregado. O erro obtido nesta camada é propagado para que os pesos dos filtros das camadas convolucionais sejam ajustados também. Em seguida, nas unidades de saída, utiliza-se uma função *softmax* para se obter a probabilidade de dada uma imagem de entrada pertencer a uma classe na saída.

Uma ilustração das camadas supracitadas é apresentada na Figura 2.



60

Figura 2: Ilustração geral sobre as camadas que formam uma rede convolutiva.

FUNDAMENTAÇÃO TEÓRICA

Na literatura, observamos que o desenvolvimento de sistemas de reconhecimento espécies florestais tem atraído o interesse de vários centros de pesquisa nos últimos anos. Entretanto, os requisitos em termos de pré-processos aplicados a madeira antes que esta seja processada computacionalmente faz com que estas abordagens sejam de difícil aplicação na prática. Mas recentemente, em [4] foi proposto um sistema de reconhecimento capaz de reconhecer espécies de madeiras a partir de câmeras de baixo custo, demonstrando que é possível resolver este problema. Técnicas de classificação de textura têm sido explorado

por vários autores nos últimos anos, assim como: o uso de filtros Gabor e co-ocorrência matrizes (GLCM), Padrões binários (LBP) para extrair recursos relevantes de as imagens, etc. Recentemente, modelos de aprendizagem profunda (Deep Learning) estão sendo utilizados, como o estado da arte em muitos domínios, como revisto por Bengio em [1], Além de melhorar a precisão em diferentes padrões problemas de reconhecimento, essa técnica move o aprendizado de máquina para o descoberta de múltiplos níveis de representação.

PROTOCOLO EXPERIMENTAL

A base de imagens original de 2240 amostras coloridas foram convertidas em nível de cinza e dividida em 3 conjuntos independentes: treinamento 10 amostras por classe, validação e teste 6 e 4 respectivamente. Para diminuir a complexidade da rede as imagens originais foram reduzidas de 1024x768 para 640x640. As bases de treinamento e validação foram utilizadas para a coleta de resultados e análises dependendo dos diferentes parâmetros aplicados. Ao final, o modelo que apresentou os melhores resultados sobre o conjunto de validação foi utilizada para a obtenção dos resultados finais sobre o conjunto de teste.

O conjunto de imagens citado acima apresenta uma diversidade signi cativa de espécies orestais. Das 112 espécies, 37 são do tipo Conífera e 75 são do tipo Folhosa. Estas 112 espécies também podem ser divididas em 85 gêneros e 30 famílias. Deste modo, o uso desta base de imagens nos permite desenvolver um sistema de reconhecimento sem preocupar-se, neste momento, com a captura de imagens. Todavia, planeja-se contribuir para o enriquecimento desta base por meio da inclusão de novas espécies futuramente, trabalhando em conjunto com o departamento de Engenharia Ambiental da UFPR.

Na Figura 3 apresentamos alguns exemplos de imagens extraídas da base de Martins et al [2012]. Nestes exemplos é possível notar que existe uma certa semelhança entre os padrões de textura de espécies pertencentes ao mesmo tipo (Conífera ou Folhosa). Mas é possível notar também uma certa distinção entre os padrões de tipos diferentes.



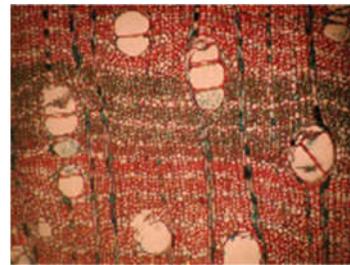
(a) *Ginkgo biloba*



(b) *Cedrus sp*



(c) *Anadenanthera colubrine*



(d) *Cariniana estrellensis*

Figura 3: Amostras de espécies florestais que constituem a base de imagens.

RESULTADOS

Na Tabela 1 tem-se uma lista de resultados envolvendo métodos de extração e com diferentes técnicas de conjunto de classificadores. Primeiramente, considerou-se as técnicas de extração GLCM, LPQ, LBP utilizando um sistema de divisão da imagens em quadrantes e combinando seus resultados para a decisão.

Logo após, utilizou-se técnicas de geração de conjuntos de classificadores (Bagging e Random Subspaces) com cada uma das técnicas de extração de características mencionadas. Finalmente, aplicou-se a técnicas de aprendizagem profunda considerando a metodologia apresentada na Seção 2. Foram testadas diversas configurações, entretanto experimentos com os parâmetros para a arquitetura da rede neural convolutiva seguir produziram os melhores resultados: 1) Camada de entrada: Entradas de 640x640 pixels e 1 canal (nível de cinza); 2) Duas combinações de camadas convolucionais e agrupadas: cada camada convolucional tem 32 filtros com tamanho $\times 5$ - e stride definido como 1, e as camadas de pool consistem de

janelas com tamanho 3×3 e stride 1; 3) Camada conectada localmente: - filtros de tamanho 128; 4) Camada de saída totalmente conectada (112). Com base nos resultados, observa-se que as técnicas de Random Subspaces e Deep Learning alcançaram os melhores resultados. Entretanto, além de obter a melhor performance, a técnica de Deep Learning (CNN) utiliza apenas um classificador em comparação com a utilização de múltiplos classificadores.

Técnica de Extração de Características	1-NN
LPQ - Uniform window- (16)	82,96 (1,26)
LPQ - Gaussian window - (16)	81,92 (1,46)
LPQU+LPQG+LPQB	87.05 (0.95)
Bagging LPQ (16 membros)	89.90 (0.70)
Random Subspaces (16 membros)	90.46 (0.81)
CNN	91.27 (0.45)

Tabela 1: Resultados obtidos em comparação com resultados anteriores

CONCLUSÃO

Os resultados demonstraram que a utilização de redes neurais convolutivas (Deep learning) produziu resultados superiores para a classificação, mesmos nos casos comparados com conjunto de classificadores, demonstrando que a técnica é eficiente para a extração e classificação das espécies florestais. Porém, cabe ressaltar que o tempo de treinamento e ajustes de parâmetros é lento.

PRINCIPAIS REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Y. Bengio and A. Courville, “Deep learning of representations,” in Handbook on Neural Information Processing. Springer, 2013, pp. 1–28.
- [2] A. Krizhevsky and G. Hinton, “Learning multiple layers of features from tiny images,” Master’s thesis, Department of Computer Science, University of Toronto, 2009.
- [3] Kapp, M. N.; Bloor, R. ; Cavalin, P. R. ; Oliveira, L. E. S. . Automatic Forest Species Recognition based on Multiple Feature Sets. In: International Joint Conference on Neural Networks (IJCNN2014) Beijing-China.
- [4] P. L. Paula Filho, L. S. Oliveira, A. Britto Jr e R. Sabourin, Forest Species Recognition using Color-based Features, October, 2010.