**Classificação de imagens Macroscópicas de espécies de madeiras no estado de São Paulo “NativeOn”**

* T. Thales Fernando Ferreira Alonso, A. F Fernando Almeida, (Sérgio Ricardo Borges Júnior: orientador).

E-mail:

alonso.thales@gmail.com; [fernandofatec2017@gmail.com](mailto:fernandofatec2017@gmail.com).

# Resumo: A exploração e corte ilegal da floresta tropical amazônica, cresceu muito nos últimos anos. Sempre voltada para prática de extração ilegal de madeira. Estima-se que, anualmente, 24 milhões de m³ de madeiras sejam extraídos da Floresta Amazônica, dos quais, aproximadamente 15%, são absorvidos pelo mercado do estado de São Paulo. Estes dados elevam o estado de São Paulo como o maior consumidor mundial de madeira nativa da floresta Amazônica. A uma grande necessidade de fiscalização por parte de órgãos legais do estado, é necessário estancar drasticamente está péssima estatística. Neste contexto foi feita uma parceria com a Polícia Ambiental do estado de São Paulo, a fim de desenvolver um método, onde será automatizado o processo de fiscalização de madeiras ilegais que entram pelas principais rotas do estado.

**Palavras-chave:** Inteligência Artificial. Identificação de madeiras. Descritores de Texturas. Classificação de imagens.

***Abstract:*** *The illegal logging and harvesting of the Amazon rainforest has grown significantly in recent years. Always focused on practicing illegal logging. It is estimated that, annually, 24 million m³ of wood is extruded from the Amazon Forest, of which, approximately 15%, is absorbed by the market of the state of São Paulo. These data elevate the state of São Paulo as the largest consumer of wood native to the Amazon rainforest. A great need for oversight by legal bodies of the state. It is necessary to stake drastically this terrible statistic. In this context, a partnership was made with the Environmental Police of the state of São Paulo, in order to develop a system, which will automate the process of inspection of illegal timber entering the main routes of the state.*

***Keywords:*** *Artificial Intelligence. Identification of wood. Textures Descriptors. Classification of images.*

***SUMÁRIO***

1. Introdução 3

2. Justificativa 3

3. Objetivo(s) 4

3.1 Objetivo Geral: 4

3.2 Objetivos específicos: 4

4. Fundamentação Teórica 4

5. Trabalhos Similares 4

6. Machine Learning 5

6.1 Redes Neurais 6

6.2 Deep Learning 6

6.3 Probabilistic Machine Learning 6

6.4 Aprendizagem Supervisionada 7

7 Metodologia 7

7.1 Algoritmo KNN / MLP Classifier 8

8 Processamento Digital de Imagem 8

8.1 Fundamentos de Imagens Digitais 9

8.2 Etapas do Processamento de Imagem 13

8.3 Histograma da Imagem 16

8.4 Realce de Imagens 17

8.5 Segmentação Orientada a Regiões 19

9 Referências 22

# Introdução

A grande devastação dos recursos naturais dos biomas florestais brasileiros, impulsionada pela exploração ilegal da madeira, tem resultado em números alarmantes de espécies de madeiras em extinção em nosso país. A madeira e seus produtos tem importância fundamental na vida de milhões de seres humanos, sendo utilizada desde o começo da civilização, seja para transporte, habitação, fabricação de armas, produção de papel e fonte de energia.

Estima-se que, anualmente, 24 milhões de m3 de madeira em toras sejam cortadas de árvores da Floresta Amazônica (IBAMA, 2011), cerca de 15% são utilizados pelo mercado do estado de São Paulo.

O estado de São Paulo constitui-se o maior mercado consumidor mundial de madeira de árvores tropicais da Amazônia, considerando 3.600.000 m3 de madeira transportados, distribuídos e armazenados pelos mais diferentes segmentos do País.

Há uma grande defasagem na fiscalização de madeiras no País. A fiscalização de madeiras ilegais no país é antiquada e lenta, é necessário um sistema que automatize esse processo, que hoje é muito lento e restrito para uma minoria de pessoas. O presente trabalho visa auxiliar o órgão competente para automatizar o processo de identificação de uma espécie de madeira, com esse processo mais rápido e confiável, obtêm-se melhores resultados para distinguir o tipo de madeira que um mesmo transeunte ou “Empresa” está transportando em nosso estado.

# Justificativa

Com o crescimento exponencial do desmatamento da Amazônia e a defasada fiscalização de madeira ilegal no estado, houve a necessidade de criar um método para auxiliar na fiscalização de madeiras no estado de São Paulo. Atualmente o sistema de fiscalização de madeira no País é muito engessado e não está disponível para todas as pessoas que precisam utilizar em seu dia a dia. O presente trabalho visa facilitar a fiscalização da Policia Ambiental do Estado de São Paulo, que é o órgão responsável para fiscalizar madeiras no País. A ideia é que seja um sistema intuitivo e de fácil utilização, com precisão em seus apontamentos para classificar espécies de madeiras. Dessa forma os órgãos competentes terão um auxilio, rápido e preciso para utilizar em seu dia a dia.

A parte mais trabalhosa desse trabalho é conseguir classificar as imagens de madeiras com precisão, então nosso foco nesse presente trabalho é conseguir classificar imagens de madeiras por espécies, o desenvolvimento do sistema para classificar imagens de madeiras, ficará como trabalho futuro.

# Objetivo(s)

# Objetivo Geral:

Identificar 4 espécies de madeiras comercializadas nos municípios do Estado de São Paulo por meio de técnicas de aprendizado de máquina e descritores de texturas.

# Objetivos específicos:

Definir principais critérios para segmentar imagem;

Definir Padrões com base em descritores de texturas;

Classificação de imagens de madeiras por espécies;

Compreender o conceito de algoritmos de processamento de imagens;

Definir Padrões para classificação;

Segmentar imagem;

# Fundamentação Teórica

A fundamentação teórica desde presente trabalho está embasada em conceitos de inteligência artificial com aprendizado de máquina. Utilizaremos conceitos de processamento de imagens e descritores de texturas.

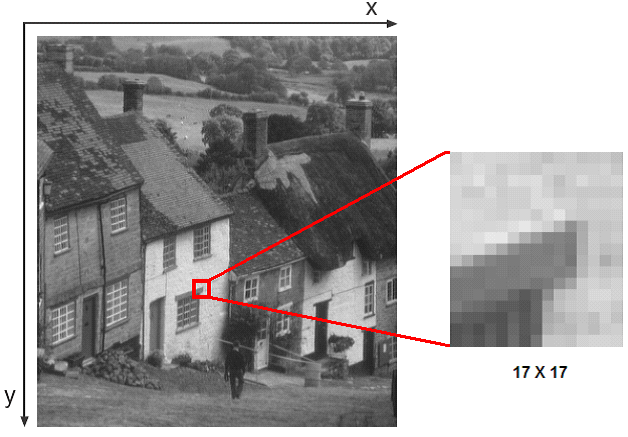
# 4.1 Processamento Digital de Imagem

Os primeiros registros de estudos em processamento de imagens digitais datam da década de 60, impulsionados pelos projetos de pesquisa criados pela NASA, nos EUA. Em seguida, surgiriam mais atividades que necessitariam dos adventos proporcionados por esta área de concentração de conhecimento. Dentre elas estavam a medicina, a microscopia, a meteorologia, o televisionamento e a indústria (GONZALEZ; WOODS, 1992). As imagens são produzidas por uma variedade de dispositivos físicos, tais como: câmeras e vídeo-câmeras; equipamentos de radiografia; microscópios eletrônicos, magnéticos e de força atômica; radares; equipamento de ultrassom. A produção e utilização de imagens podem ter diversos objetivos, que vão do puro entretenimento até aplicações militares, médicas ou tecnológicas (ESQUEF, 2002).

O processamento digital de imagem (PDI) é uma área da eletrônica (teoria de sinais) em que imagens são convertidas em matrizes de números inteiros, sendo que cada elemento desta matriz é composto por um elemento fundamental: o pixel (abreviação, do Inglês: *Picture element*). A partir desta matriz de pixels, que representa a imagem, diversos tipos de processamento digital podem ser implementados por algoritmos computacionais. A aplicação destes algoritmos realiza as transformações necessárias para que se possa, por exemplo, obter uma imagem com os realces pretendidos ou extrair atributos e informações pertinentes (GONZALEZ;WOODS, 1992; COTRIM, 2007).

# Fundamentos de Imagens Digitais

O computador não pode guardar em sua memória ou modificar em seus circuitos qualquer imagem na forma analógica, que é como o olho humano a enxerga, porque o computador trabalha com números discretos, não podendo representar diretamente tons de cinza ou cores contínuas. Para que o computador possa operar com imagens, primeiro elas precisam ser convertidas para uma grande lista de números (COTRIM, 2007). Uma imagem monocromática1 é uma função bidimensional f (x; y) da intensidade luminosa, onde x e y denotam coordenadas espaciais, que por convenção: x = [1; 2;: : : ; M] e y = [1; 2; : : : ; N]. O valor de f no ponto (x; y) é proporcional ao brilho (ou nível de cinza) da imagem nesse ponto, como observado na Figura 1, cuja região em destaque apresenta os pixels e os níveis de luminância de cada um deles (ESQUEF, 2002).

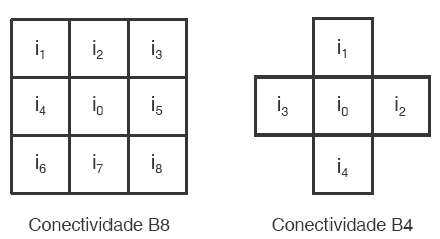


**Figura 1** Imagem monocromática com destaque para uma região de 17 \_ 17 pixels.

Fonte: Esquef (2002), modificado pelo autor.

Com isso, um pixel é o elemento básico em uma imagem, normalmente representada em forma retangular ou quadrada, e suas dimensões são finitas na representação de uma imagem digital. Frequentemente a organização de uma imagem sob a forma de uma matriz de pixels é feita em uma simetria quadrada, cujo objetivo é facilitar a implementação eletrônica, seja dos sistemas de aquisição, seja dos sistemas de visualização de imagens (MASCARENHAS; VELASCO, 1989).

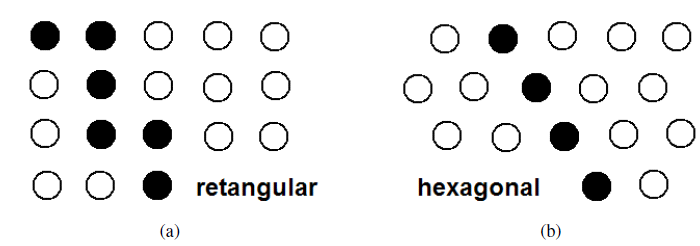
Entretanto, um pixel não apresenta as mesmas propriedades em todas as direções (i.e. ele é anisotrópico) e esta propriedade faz um pixel ter quatro vizinhos de borda e quatro vizinhos de diagonal (cfe. Figura 2), sendo que esse tipo de organização provoca o aparecimento de dois problemas importantes nas técnicas de PDI (ESQUEF, 2002). O primeiro consiste em determinar qual tipo de conectividade será utilizada: B4 (vizinhos de borda) ou B8 (vizinhos de borda e os de diagonal). Já o segundo está relacionado com as distâncias entre um ponto e seus vizinhos, determinando que um ponto tenha distância igual a 1 para vizinhos de borda e p 2 para aqueles na diagonal.



**Figura 2**: Tipos de conectividade dos pixels vizinhos ao pixel central.

Fonte: Esquef (2002).

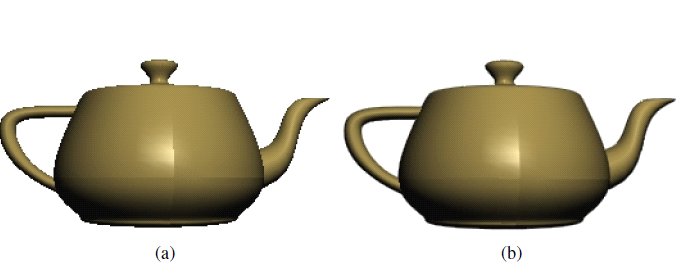
Sendo assim, para que uma imagem possa ser manipulada por computador, segundo Mascarenhas e Velasco (1989), é necessário que essa imagem seja discreteada ou digitalizada, tanto espacialmente quanto em sua amplitude. Com isso, uma imagem pode ser discreteada através de um processo que envolve dois passos, a saber: (i) amostragem, no qual são definidos os valores para as coordenadas x e y da matriz de pontos, e (ii) quantização, que consiste em escolher um valor múltiplo relacionado à banda de frequência da intensidade para cada ponto da imagem. Então, amostrar uma imagem significa definir, para uma imagem real, as características da matriz sobre a qual esta será representada. Essas características são comumente reconhecidas como resolução da imagem e estão diretamente relacionadas com a qualidade da imagem. Evidentemente, dependendo das características desta matriz, ou seja, tamanho e dimensões dos pixels, mais fiel será a representação discreta da imagem. A Figura 3 ilustra dois tipos de amostragem para representação de uma reta (EFFORD, 2000).



**Figura 3**: Diferença na geometria da amostragem: (a) retangular; (b) hexagonal.

Fonte: E\_ord (2000), modificado pelo autor.

Uma amostragem insatisfatória gerará uma reprodução de baixa qualidade, sendo o problema de aliasing2 um problema comum nesses casos, efeito esse caracterizado pela aparência serrilhada das bordas dos objetos na imagem (cfe. Figura 4 (a)). Uma solução para eliminar esse problema é utilizar um filtro denominado anti-aliasing (ver Figura 4 (b)), que diminui os serrilhados nas silhuetas dos objetos através de um conjunto de cálculos e procedimentos suavizadores da transição visual entre os modelos exibidos e o ambiente ao seu redor. Entretanto, isso está associado a um alto custo computacional (EFFORD, 2000).

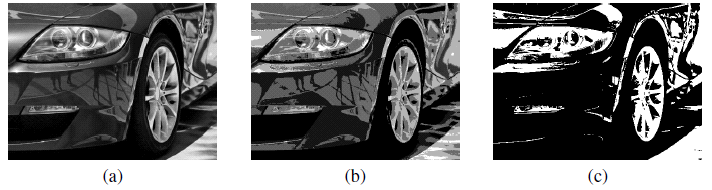


**Figura 4**: Amostragem com: (a) aliasing; (b) anti-aliasing.

Já os métodos de quantização buscam, em geral, otimizar os erros entre a imagem real e a imagem quantizada, ou seja, procuram definir variações de tons de cinza que representem, de forma mais fiel possível, os níveis de cinza de uma imagem real.

Cada tom é representado por um código binário que é formulado atendendo às necessidades de minimização da quantidade de bits, visando seu armazenamento ou sua transmissão (GONZALEZ; WOODS, 1992).

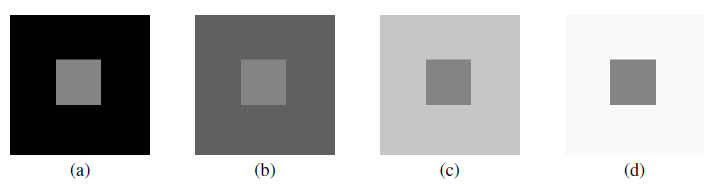
Erros como os de “falsos contornos” podem ocorrer quando se usa um reduzido número de tons de cinza em uma quantização (ver Figura 5 (b) e (c)). Este fenômeno caracterizasse pela transição de tons de cinza mais frequentes em áreas de pouca variação. De uma forma geral, pode-se dizer que, dispondo de uma dada quantidade de bits, a escolha do nível de quantização e da amostragem dependerá do tipo de imagem a ser discretizada e dos critérios de avaliação subjetivos (MASCARENHAS; VELASCO, 1989)



**Figura 5:** Quantização monocromática: (a) 256 níveis; (b) 16 níveis; (c) 2 níveis.

Não existe uma regra exata para determinado tipo de imagem, no entanto, é possível concluir que, para imagens com variação suave de tons é necessária uma maior quantidade de níveis de quantização, não sendo relevante a amostragem. Já para imagens com grande detalhamento é conveniente o uso de uma amostragem densa, não sendo necessária uma quantização com muitos níveis de cinza (GONZALEZ; WOODS, 1992).

Quanto ao contraste de uma imagem, segundo Gonzalez e Woods (1992), pode ser definido como a razão dos níveis de cinza do objeto e do fundo (diferença local de luminância). Dessa afirmação deriva-se o brilho aparente de um objeto, o qual depende fortemente da intensidade do fundo, sendo que qualquer alteração no contraste pode gerar uma melhor definição dos objetos. A Figura 6 demonstra para um objeto de mesmo nível de cinza diferentes fundos, tornando o objeto mais ou menos visível.



**Figura 6**: Exemplo de variação de contraste.

Fonte: Gonzalez e Woods (1992).

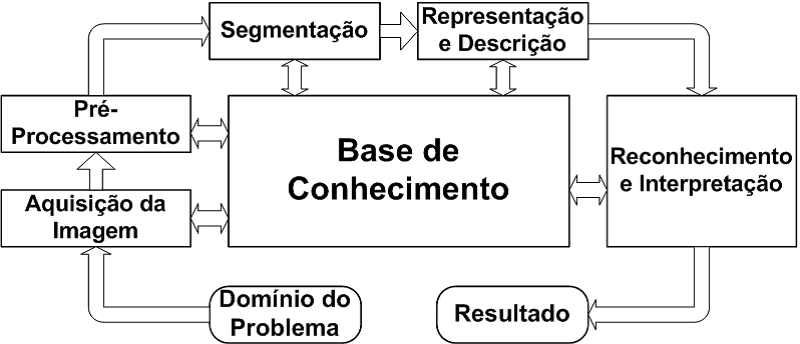
Imagens que apresentam variação de valores em mais de uma banda de frequência necessitam ser representadas por mais de uma função do tipo f (x; y). As imagens coloridas são um exemplo, pois apresentam uma função de intensidade para cada cor primária (RGB).

A sobreposição destas três imagens, que individualmente são monocromáticas, compõe uma imagem colorida (GONZALEZ; WOODS, 1992).

O tipo de imagem e a qualidade desta dependem necessariamente de como a imagem foi obtida, e os sistemas responsáveis pela obtenção ou captura de imagens são denominados sistemas de imageamento. Estes são sistemas físicos destinados a produzir ou capturar imagens e, apesar de diferirem na aparência, apresentam várias características comuns. Dentre os mais comuns estão a máquina fotográfica, a câmera de vídeo, os sensores a bordo de um satélite e os equipamentos de raio-X (MASCARENHAS; VELASCO, 1989).

# Etapas do Processamento de Imagem

Alguns autores possuem conceitos diferentes quanto à taxonomia do processo de um sistema PDI. Entretanto, seja uma classificação mais detalhada ou não do processo, todos seguem uma mesma linha que, neste caso, segue a classificação segundo Gonzalez e Woods (1992), a saber: formação e aquisição da imagem, pré-processamento, segmentação, representação e descrição e reconhecimento e interpretação, como ilustra a Figura 7. A seguir, serão apresentadas breves descrições de cada uma delas.



**Figura 7**: Etapas do processamento digital de imagens.

Fonte: Gonzalez e Woods (1992)

O processo de aquisição da imagem consiste no primeiro passo do processo e necessita dispor de um dispositivo físico que deve ser sensível ao espectro de energia eletromagnético, como por exemplo, ao espectro de raio-X, luz ultravioleta, luz visível ou luz infravermelha. Este dispositivo transdutor deve produzir em sua saída um sinal elétrico proporcional ao nível de energia percebido. A partir disso, a etapa de digitalização da imagem (omitida na ilustração anterior 7) necessita dispor de um dispositivo, chamado digitalizador, que converte o sinal elétrico analógico produzido na saída do sensor em um sinal digital (GONZALEZ; WOODS, 1992; ESQUEF, 2002).

A fase de pré-processamento é a que se utilizam técnicas para melhorar a qualidade da imagem de forma a aumentar as chances de sucesso nos processos seguintes. Estas técnicas envolvem duas categorias principais: métodos que operam no domínio espacial e métodos que operam no domínio da frequência (GONZALEZ; WOODS, 1992; ESQUEF, 2002). Técnicas de processamento no domínio espacial baseiam-se em filtros que manipulam o plano da imagem, enquanto as técnicas de processamento no domínio da frequência baseiam-se em filtros que agem sobre o espectro da imagem. É comum para realçar determinadas características de uma imagem combinar vários métodos que estejam baseados nestas duas categorias (GONZALEZ; WOODS, 1992; ESQUEF, 2002).

O estágio de segmentação é aquele em que se divide uma imagem de entrada em partes ou objetos constituintes que se diferenciam entre si. Essas denominações “objeto” e “partes” possuem uma conotação bastante subjetiva, podendo se referir a grupos de pixels que formam determinadas regiões na imagem, sem que representem literalmente um objeto presente na imagem (GONZALEZ; WOODS, 1992).

A segmentação é considerada, dentre todas as etapas do PDI, a mais crítica do tratamento da informação. É nesta etapa que são definidas as regiões de interesse para processamento e análise posteriores. Como consequência, quaisquer erros ou distorções presentes nesta etapa refletem-se nas demais, de forma a produzir ao final do processo resultados não desejados, podendo contribuir de forma negativa para a eficiência de todo o processamento, premissa essa também verdadeira para o caso inverso (ESQUEF, 2002).

Deve-se ressaltar que não existe um modelo formal para a segmentação de imagens.

A segmentação é um processo empírico e adaptativo, procurando sempre se adequar às características particulares de cada tipo de imagem e seus objetivos. Apesar da grande diversidade de técnicas de segmentação de imagens, ainda assim, atualmente existe um grande interesse no estudo e desenvolvimento de novas técnicas (ESQUEF, 2002). Diferentemente de Gonzalez e Woods (1992), alguns autores consideram uma etapa intitulada pós-processamento geralmente sucedendo a fase de segmentação, consistindo em corrigir os principais defeitos ou imperfeições da segmentação. A escolha de uma representação é apenas parte da solução para transformar os dados iniciais numa forma adequada para o subsequente processamento computacional. Um método para descrever os dados também deve ser especificado, de forma que as características de interesse sejam enfatizadas. O processo de descrição, também chamado seleção de característica ou extração de atributos, consiste em obter informações numéricas dos objetos contidos na imagem, podendo ser a etapa final de um sistema de PDI, caso este seja seu objetivo (GONZALEZ; WOODS, 1992; ESQUEF, 2002).

Já o **reconhecimento** é o processo que atribui um rótulo a um objeto, baseado na informação fornecida pelo seu descritor. A interpretação envolve a atribuição de significados a um conjunto de objetos reconhecidos. Um exemplo para essa etapa é o reconhecimento de um caractere, a letra “c” por exemplo, contido em uma imagem. Para identificação do *caracter* há a associação dos descritores para aquele *caracter* com o rótulo [c]. A partir daí, a interpretação procura atribuir um significado a um conjunto de entidades rotuladas (GONZALEZ; WOODS, 1992).

O conhecimento sobre o domínio do problema está codificado em um sistema de PDI na forma de uma base de conhecimento. E, esse conhecimento pode ser tão simples quanto o detalhamento de regiões de uma imagem nas quais se sabe que a informação de interesse pode ser localizada, limitando, assim, a busca que precisaria ser feita para obter aquela informação. Em contrapartida, a base de conhecimento pode ser bastante complexa, tal como uma base de imagens contendo imagens de satélite de alta resolução de uma região em conexões com aplicações de detecção de mudanças (GONZALEZ; WOODS, 1992; ESQUEF, 2002).

Além de guiar a operação de cada módulo de processamento, a base de conhecimento também controla a interação entre os módulos. Essa distinção é demonstrada na Figura 7 através das flechas duplas entre os módulos de processamento e a base de conhecimento, que, ao contrário das flechas simples, ligam apenas os módulos de processamento. Essa representação indica que a comunicação entre os módulos de processamento baseia-se normalmente em conhecimento prévio da natureza esperada do resultado (GONZALEZ; WOODS, 1992).

# Realce de Imagens

A manipulação de imagens inclui duas grandes classes de transformações, a Saber: transformações radiométricas, em que se alteram apenas os valores dos níveis de cinza, e transformações geométricas, em que a geometria da imagem é alterada, mantendo-se, ao máximo possível, os valores dos níveis de cinza (GONZALEZ; WOODS, 1992).

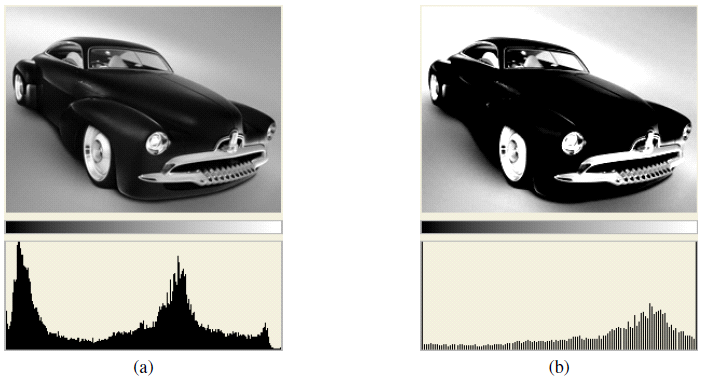
Podem-se distinguir dois grupos de transformações radiométricas: a restauração, que visa corrigir distorções sofridas pela imagem, geradas normalmente pela limitação dos sistemas de imageamento, e a técnica que será abordada nesta Seção intitulada realce, que visa enfatizar alguma característica de interesse na imagem. Em determinados casos estas duas transformações produzem os mesmos resultados (JAIN, 1989 apud ANTUNES, 1999).

O **realce** de imagem tem por objetivo acentuar certos aspectos da imagem para subsequente análise ou visualização. No realce de imagem não são acrescidas informações a ela, e sim enfatizadas determinadas características da imagem. Um exemplo de realce é a variação de contraste, que transforma o valor de cinza de determinados pixels em valores diferentes, ressaltando determinados aspectos da imagem (GONZALEZ; WOODS, 1992 apud ANTUNES, 1999).

As técnicas de realce de imagem, segundo Gonzalez e Woods (1992), apresentam duas Grandes categorias: (i) os métodos **ponto a ponto** e (ii) os métodos de **vizinhança.**

Os métodos ponto a ponto caracterizam-se por modificar os níveis de cinza de uma imagem preservando os contornos. São operações aplicadas sobre cada pixel, sendo um exemplo deste método a equalização de histograma (GONZALEZ; WOODS, 1992).

Lembre-se que um histograma é uma tabela que indica para cada nível de cinza a quantidade de pixels existente na imagem, apresentando uma frequência ou densidade maior de pixels em um determinado nível de cinza que em outro. Portanto, equalizar um histograma, segundo Gonzalez e Woods (1992), significa gerar uma uniformização de densidade para os níveis de cinza da imagem, resultando em maior contraste e nitidez da imagem, ressaltando, assim, os objetos nela contidos, conforme ilustra a Figura 8 (b).



**Figura 8** Equalização de histograma: (a) imagem original; (b) imagem equalizada.

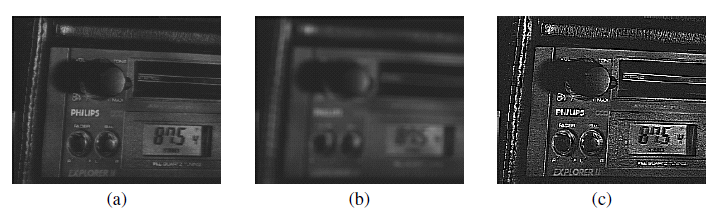
Já os métodos de **vizinhança** são aplicados sobre uma determinada região da imagem para enfatizar determinados aspectos desta (GONZALEZ; WOODS, 1992). Uma das aplicações destas técnicas de vizinhança é a eliminação de ruídos através da suavização de imagens. E, um exemplo interessante para representar este método é a técnica de filtragem.

Na técnica de filtragem, o valor de nível de cinza de uma imagem depende não apenas do nível de cinza original, mas dos valores de níveis de cinza de sua vizinhança. O tamanho desta vizinhança não é limitado, sendo que normalmente os vizinhos mais próximos têm maior influência sobre os mais distantes (JAIN, 1989 apud ANTUNES, 1999).

Os filtros, de modo geral, são projetados e implementados como ferramentas para realizar realce de imagens, cuja atividade consiste no incremento da qualidade geral da imagem, podendo ser interessante para fins de visualização, ou ainda, para ressaltar determinados objetos e características da imagem (JAIN, 1989 apud ANTUNES, 1999). Existem duas classes importantes de filtros amplamente utilizadas: (i) filtros de passa-baixa e (ii) filtros de passa-alta.

Um filtro passa-baixa tem valores próximos a zero para as altas frequências, portanto, a imagem apresenta-se “suavizada”. Isto ocorre porque as altas frequências correspondem às transições abruptas. Este filtro tem ainda como efeito a eliminação de ruídos da imagem, mas apresenta defeitos ao tornar a imagem menos nítida e definida (JAIN, 1989). A Figura 9 (b) ilustra um filtro passa-baixa denominada filtro da média.

Um filtro passa-alta tem valores próximos a zero para as baixas frequências, resultando em um efeito de “agudização” da imagem, ou seja, as transições entre diferentes regiões da imagem tornam-se mais nítidas. Entretanto, ele enfatiza o ruído que possa existir na imagem (JAIN, 1989). A Figura 9 (c) ilustra um filtro passa-alta denominado filtro laplaciano + original, usado para detectar bordas mantendo a informação original.



**Figura 9**: Filtragem: (a) imagem original; (b) filtro passa-baixa; (c) filtro passa-alta.

O processo de filtragem é feito utilizando matrizes denominadas máscaras, as quais são aplicadas sobre a imagem. Esse processo também é conhecido como convolução. Assim sendo, convolução é o processo de calcular a intensidade de um determinado pixel em função da intensidade de seus vizinhos, o cálculo é baseado em ponderação, isto é, utilizam-se pesos diferentes para pixels vizinhos diferentes (GONZALEZ; WOODS, 1992).

# Segmentação Orientada a Regiões

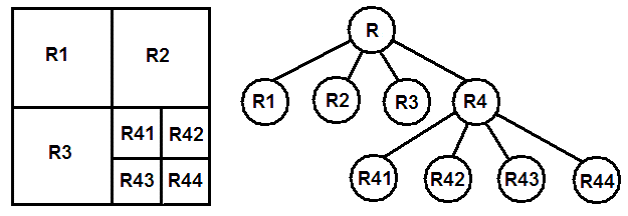
O objetivo da segmentação de imagens é obter, a partir de uma imagem digitalizada pré-processada, um conjunto de primitivas ou segmentos significativos, que contenham informações semânticas relativas à imagem de origem. Normalmente, as primitivas usadas são contornos e regiões. Os diferentes tratamentos utilizados levam à noção de contornos, regiões ou mista. Entretanto, será abordada apenas a segmentação por regiões (ANTUNES, 1999).

A detecção de regiões em uma imagem pode ser feita com dois objetivos: extrair uma determinada região ou dividir a imagem num conjunto de regiões distintas. Uma região de uma imagem é definida como um conjunto de pontos ligados em que, a partir de qualquer ponto da região pode-se chegar a qualquer outro ponto por um caminho completamente contido nessa região (GONZALEZ; WOODS, 1992 apud ANTUNES, 1999).

Essas regiões normalmente apresentam alguma característica homogênea, que comumente é a continuidade do nível de cinza de seus pixels (GONZALEZ; WOODS, 1992 apud ANTUNES, 1999). Os métodos de **divisão e fusão** e **limiarização** podem ser citados como diferentes estratégias nesse contexto de segmentação por região.

O método de **divisão e fusão** consiste em agrupar os pixels até que se tenham regiões sobre as quais exista uma homogeneidade H. Para isto são executadas sucessivas divisões e fusões sobre a imagem. Este método é normalmente associado a uma estrutura de dados quadtree7, que permite decompor e agrupar partes de uma imagem (GONZALEZ; WOODS,1992).

Construir essa estrutura quadtree significa dividir uma região em quatro quadrantes, sendo que cada um pode ser dividido novamente em quatro subquadrantes, e assim sucessivamente, até que seja encontrado o fator H. A Figura 10 ilustra uma segmentação quadtree (esq.) e seu grafo representando seus particionamentos (dir.) (GONZALEZ; WOODS, 1992).

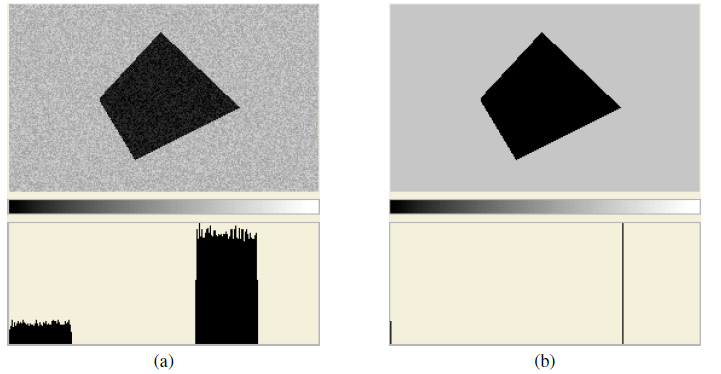


**Figura 10**: Imagem segmentada usando quadtree (esq.) e grafo da segmentação (dir.).

Fonte: Gonzalez e Woods (1992), modificado pelo autor.

7 Uma quadtree é uma estrutura de dados utilizada para codificar imagens.

O princípio da **limiarização** consiste em separar as regiões de uma imagem quando esta apresenta duas classes, o fundo e o objeto, utilizando-se do histograma para isto. Caso se tenha uma imagem com fundo e objeto bem distintos (níveis de cinza com uma variação bem definida), o histograma dessa imagem terá dois picos bem distintos separados por um vale de valores relativamente baixos (GONZALEZ; WOODS, 1992). Pode-se observar na Figura 11 uma imagem com ruídos (a), bem como o resultado após a aplicação da limiarização (b).



**Figura 11**: Imagem e histograma: (a) com ruído; (b) pós limiarização.

**4.6 Definições de Inteligência Artificial**

Em primeiro lugar, a preocupação da IA era fazer a máquina pensar de forma humana e estava preocupada com processos de pensamento e raciocínio. Além disso, mede o sucesso em termos de desempenho humano. Nesse sentido, Bellman (1978) delineou-o como a mecanização de atividades ligadas à compreensão humana; e práticas como aprendizagem, decisão lógica e gerenciado. Ao longo da década de 1980, os especialistas definiram a AI como a capacidade de fazer um computador pensar literalmente e racionalmente, destacando a própria inteligência e as faculdades mentais (CHARNIAK; McDERMOTT, 1985; HAUGELAND, 1985; MOIGNE, 1988). Haddock e O'keefe (1990) concluíram que a AI é apenas uma inovação empolgante que ajudará a modelar e a examinar melhor os sistemas de fabricação, a fim de colocar o método de investigação sob o controle de pessoas com os problemas. Como observado, em menos de uma década, as definições de AI passaram da imaginação à modelagem para a atuação, como pode ser visto nas opiniões de Kurzweil (1990); Rich and Knight (1990); e Schalkoff (1990), que caracterizaram a Inteligência Artificial como a capacidade de raciocinar e agir de forma humana, e realizar atividades da maneira correta. Em breve, os autores começaram a comparar AI com o comportamento humano, descrevendo-o como a automação da inteligência e o emulador de comportamento de inteligência (DE SILVA, 1994; LUGER, 1995; LIBERATORE, 1997).

Aprendizagem de máquinas - ML é parte da Inteligência Artificial que auxilia os computadores com a capacidade de aprender sem serem claramente programados, ou seja, sem orientação humana. A Aprendizagem de Máquinas concentra-se no avanço de programas de computador, que são capazes de instruir-se a desenvolver e mudar quando expostos a novas informações (MEYER et al., 2016).

Os algoritmos de inteligência artificial tornaram-se um mecanismo promissor para lidar com alguns problemas de investigação extrema. O aprendizado de máquina, como uma subárea expansiva da AI, está preocupado com cálculos e sistemas que permitem que os computadores aprendam. O aprendizado da máquina abrange os principais espaços, por exemplo, aplicações de mineração e programação de informação (ALAVI; GANDOMI; LARY, 2016).

Conforme descrito por Franco e Bacardit (2016), o procedimento de aprendizado automático é como o da mineração de dados. Ambos digitalizam informações para procurar exemplos padrão. No entanto, ao invés de extrair informações para a compreensão humana, o aprendizado de máquina utiliza essa informação para melhorar a própria compreensão do sistema. Os programas de aprendizagem de máquinas distinguem os desenhos em dados e alteram as atividades do programa adequadamente.

De acordo com a definição de Ertugrul (2016), a ML faz procedimentos aplicando as informações necessárias para que uma máquina aprenda e se ajuste quando apresentada a novos dados. É considerado como um treinador de máquinas. Baseia-se em outras duas técnicas, examinando informações minadas, fazendo um algoritmo via Inteligência Artificial, e depois disso, atualizando o presente algoritmo adequadamente para assimilar outra tarefa.

O autor continua dizendo que o ML está equipado para resumir dados de conjuntos de dados extensivos e depois identifica e deduz padrões com um objetivo específico para aplicar esses dados a novos arranjos e atividades. Claramente, os parâmetros necessários devem ser construídos em direção ao início do método de aprendizagem da máquina para que a máquina possa descobrir avaliar e acompanhar novas informações.

Fernandes, Ramos e Rosa (2005) e Mishra (2011) afirmam que essas abordagens da Inteligência Artificial, ou Modelos de Aprendizado de Máquinas, são distribuídas, como se vê na Tabela 1:

|  |  |
| --- | --- |
| Inteligência Artificial | Características |
| Algoritmo Genético | Modelo usado para a aprendizagem de máquinas. Além disso, é um método usado por algoritmos evolutivos. O objetivo deste modelo é simular operadores genéticos específicos, como *crossover*, mutação e reprodução. |
| Lógica Fuzzy | Modo de raciocínio que usa expressões lógicas para demonstrar o grau de associação em conjuntos difusos. |
| Sistema Baseado em Regras | Sistemas que realizam comportamentos inteligentes de conhecimento humanos. |
| Programação Genética | Área dedicada à construção de programas, a fim de copiar o processo natural de genética usando a busca aleatória. |
| Razão Baseada em Casos | Representa e processa conhecimento para resolver problemas através da recuperação e adaptação de experiências, que são intituladas como casos. |
| Rede Neural | Área que concentra o processamento e armazenamento de conhecimento experimental e disponibiliza-o. O conhecimento obtido pela rede de acordo com seu ambiente através da aprendizagem e as conexões entre os neurônios. |
| Árvore de Decisão | É uma das técnicas mais bem-sucedidas para a aprendizagem de classificação supervisionada. Ele pressupõe que todos os recursos possuem domínios discretos finitos, e há um único recurso de destino chamado classificação. |
| Máquina de Vetor de Suporte | O SVM converte o espaço de entrada em um espaço de característica dimensional maior, onde os dados são separáveis. |
| Rede Neural Convolucional | Eles apresentam modelos de várias camadas totalmente capazes de serem treinados, com características de captura de mapeamentos não lineares através de entradas / saídas. |
| K-vizinhos Mais Próximos | KNN classifica imagens vizinhas para o caso no caso, de acordo com distâncias euclidianas. |

Tabela 1 - Inteligência Artificial. Fonte: Adaptado de Fernandes, Ramos e Rosa (2005) e Mishra (2011)

# Trabalhos Similares

Dois sistemas de visão artificial, que usam imagens para identificar e classificar madeiras, foram desenvolvidos recentemente em São Paulo. Um deles, chamado Neuro-Wood, contou com pesquisadores da Universidade Estadual Paulista (UNESP), campusde Itapeva, e do Instituto de Ciências Matemáticas e de Computação (ICMC) da Universidade de São Paulo (USP) em São Carlos. Ele é composto por um conjunto de câmeras (webcams), um computador e um programa que diferencia a madeira em três categorias: A (excelente), B (boa) e C (rejeitada). O outro, criado no Instituto de Física da USP em São Carlos (IFSC-USP), é um método matemático que deu origem a um software capaz de determinar a espécie de árvore da qual determinada tábua provém. As duas tecnologias se destinam principalmente aos setores madeireiro e moveleiro.

Normalmente a indústria de madeira usa especialistas que classificam a qualidade das peças por meio de inspeção visual. O processo é subjetivo e depende da qualidade do treinamento, o que torna o índice de acerto não muito alto. Estudos mostram que o nível de acurácia desse método gira em torno de 65%%. Diante desse quadro, o engenheiro mecânico Carlos de Oliveira Affonso, professor do curso de engenharia industrial madeireira da UNESP de Itapeva, o cientista da computação André Luís Debiaso Rossi, professor do curso de engenharia de produção da UNESP de Itapeva, e o engenheiro civil Fábio Henrique Antunes Vieira, professor da Faculdade de Tecnologia de Capão Bonito (SP), projetaram um equipamento para realizar a classificação de madeira de forma automática. O projeto Neuro-Wood teve o apoio do Centro de Ciências Matemáticas Aplicadas à Indústria (CeMEAI), um dos Centros de Pesquisa, Inovação e Difusão (Cepid) financiados pela FAPESP, com sede no ICMC. O sistema tem webcam, monitor e um controlador lógico programável (CPL), que é um microprocessador responsável pela interface entre o computador e os atuadores (motores elétricos ou esteiras transportadoras). O programa de computador desenvolvido usa técnicas de aprendizagem de máquina. “São semelhantes às utilizadas pelos sistemas de reconhecimento facial, só que mais simples”, conta Affonso. Foram usadas as chamadas redes neurais artificiais, técnicas computacionais que mimetizam o funcionamento do cérebro humano, aprendendo com a experiência. “Para isso, é apresentado ao computador um padrão numérico correspondente à determinada classe de objetos”, explica. “Após certo número de repetições, esses softwares conseguem identificar à qual classe o objeto pertence, mesmo que não tenha sido apresentado como exemplo.” No caso do Nero-Wood, o sistema foi “ensinado” a classificar as peças de madeira conforme sua qualidade (A, B ou C). O software foi abastecido com informações sobre os níveis de qualidade e os defeitos das tábuas, como nós e rachaduras. Em seguida, criou-se um banco com mais de 600 fotos de amostras das três qualidades. Elas foram processadas para melhorar o contraste e o brilho e ressaltar detalhes, levando em conta características, como textura e coloração.

# Machine Learning

*Machine Learning* é um subconjunto de Inteligência Artificial que permite que os sistemas aprendam e prevejam resultados sem programação explícita.

Em vez de escrever algoritmos e regras que tomem decisões diretamente, ou tentar programar um computador para "ser inteligente" usando conjuntos de regras, exceções e filtros, o *Machine Learning* ensina sistemas informáticos a tomar decisões aprendendo com grandes conjuntos de dados. Os sistemas baseados em regras rapidamente se tornam frágeis quando precisam explicar a complexidade do mundo real. A Aprendizagem Automática pode criar modelos que representam e generalizam padrões nos dados que você usa para treiná-lo, e pode usar esses modelos para interpretar e analisar novas informações.

O Aprendizado Automático da máquina é adequado para a classificação, que inclui a capacidade de reconhecer texto e objetos em imagens e vídeos, bem como encontrar padrões e associações em dados ou segmentar dados em clusters (por exemplo, encontrar grupos de clientes). O *Machine Learning* também é adepto da previsão, como o cálculo da probabilidade de eventos ou resultados de previsão. E pode ser usado para gerar dados faltantes; por exemplo, a versão mais recente do CorelDraw usa *Machine* *Learning* para interpolar o curso suave que você está tentando extrair de vários cursos difíceis que você faz com a caneta.

O coração do *Machine Learning* são os algoritmos. Alguns, como regressões, *k-means clustering* e suporte de máquinas vetoriais, estão em uso há décadas. A chave para o uso efetivo está na combinação do algoritmo certo com o seu problema.

# Redes Neurais

Uma Rede Neural é um algoritmo de Aprendizado de Máquina construído em uma rede de nós interconectados que funcionam bem para tarefas como reconhecimento de padrões.

Não é um novo algoritmo, mas a disponibilidade de grandes conjuntos de dados e processamento mais poderoso (especialmente as GPUs, que podem lidar com grandes fluxos de dados em paralelo) as tornou úteis na prática. Apesar do nome, as Redes Neurais são baseadas apenas vagamente em neurônios biológicos. Cada nó em uma rede neural possui conexões para outros nós que são acionados por entradas. Quando desencadeada, cada nó adiciona um peso à sua entrada para marcar a probabilidade de que ele faça ou não a função desse nodo. Os nós são organizados em camadas fixas através das quais os dados fluem, ao contrário do cérebro, que cria, remove e reorganiza regularmente conexões de sinapse.

# Deep Learning

O *Deep Learning* é um subconjunto do *Machine Learning* baseado em Redes Neurais profundas, que têm muitas camadas para realizar a aprendizagem em várias etapas. As chamadas *Convolutional Deep Neural Networks* geralmente realizam o reconhecimento da imagem ao processar uma hierarquia de recursos em que cada camada procura objetos mais complicados. Por exemplo, a primeira camada de uma rede profunda que reconhece raças de cães pode ser treinada para encontrar a forma do cachorro em uma imagem, a segunda camada pode observar texturas como pele e dentes, com outras camadas reconhecendo ouvidos, olhos, caudas e outras características e o nível final que distingue as diferentes raças. Já as *Recursive Deep Neural Networks* são usadas ​​para reconhecimento de fala e processamento de linguagem natural, onde a sequencia e o contexto é importante.

# Aprendizagem Supervisionada

Muitas técnicas de *Machine Learning* utilizam a aprendizagem supervisionada, em que uma função é derivada de dados de treinamento rotulados. Os desenvolvedores escolhem e rotulam um conjunto de dados de treinamento, colocam uma proporção desses dados para testes e marcam os resultados do sistema de *Machine Learning* para ajudá-lo a melhorar. O processo de treinamento pode ser complexo, e os resultados são muitas vezes probabilidades, com um sistema sendo, por exemplo, 30% mais seguro de ter reconhecido um cachorro em uma imagem, 80% de que encontrado um gato, e talvez até com 2% de certeza de ter encontrado uma bicicleta. Os *feedbacks* dos desenvolvedores ajudam o sistema definir a melhor a resposta no futuro.

É importante não treinar o sistema com precisão aos dados de treinamento; Isso é chamado de superposição e significa que o sistema não poderá generalizar para lidar com novas entradas. Se os dados mudam significativamente ao longo do tempo, os desenvolvedores precisarão treinar o sistema devido ao que alguns pesquisadores referem como "ML rot".(CIO, 2017).

**6.5 K-VIZINHOS MAIS PRÓXIMOS - KNN**

De acordo com Sánchez et al. (2016), o algoritmo de vizinhos mais próximos de K é uma técnica de aprendizado de máquina que tem sido amplamente utilizada no reconhecimento e classificação de padrões. O método KNN é um algoritmo não paramétrico que foi amplamente utilizado para tarefas de classificação. O algoritmo reúne todos os exemplos acessíveis como entrada e classifica os novos casos, como resultado, com base em uma medida de afinidade.

As técnicas KNN são uma classe de abordagens multivariadas e não paramétricas para predição contínua ou categórica. Com essas técnicas, as previsões são calculadas como combinações lineares de observações para unidades populacionais em uma amostra que é semelhante ou mais próximas em um espaço de variáveis auxiliares para unidades populacionais que requerem previsões (CHIRICI et al., 2016).

A Figura 12 traz a classificação através do KNN, e pode-se observar que duas possibilidades estão disponíveis com resultados diferentes. Se alguém considerar k = 3, a saída é distinta de k = 5, onde k representa o número de vizinhos, uma vez que a técnica reconhece a semelhança do objeto pela proximidade com outros. Em outras palavras, o KNN classifica imagens de vizinhança para o caso, de acordo com suas distâncias.



**Figura 12** - Representação dos vizinhos mais próximos. Fonte: Kim; Savarese (2012).

Conforme definido por Duda, Hart e Stock (2001), o classificador KNN explora uma matriz de dados de treinamento X, em que cada objeto faz parte de uma classe "n" fora de "N" classes concebível. Este classificador permite um objeto desconhecido xt, para a classe à qual a maior parte dos vizinhos mais próximos. Estes vizinhos são encontrados por métricas adequadas, normalmente a distância euclidiana.

Existem algumas variedades do sistema KNN, assumindo o tipo de distância utilizada ou a regra de decisão aplicada para classificação (VILLA; BOQUÉ; FERRÉ, 2008).

# 6.6 SUPORTE VECTOR MÁQUINA – SVM

As máquinas de vetor de suporte (SVMs) são um arranjo de rotinas de aprendizado regulamentadas relacionadas utilizadas para classificação e regressão (VAPNIK, 1995). Fazem parte de um grupo de classificadores lineares ambíguos. Em outras palavras, o Support Vector Machine (SVM) é um motor de predição e classificação que usa a hipótese de aprendizagem de máquinas para amplificar a precisão precoce e, consequentemente, mantendo-se longe do excesso de ajuste aos dados. As máquinas de vetor de suporte podem ser caracterizadas como estruturas, que utilizam o espaço de especulação de funções lineares em um espaço de característica dimensional elevado, preparado com um algoritmo de aprendizagem de uma hipótese simplificadora que atualiza um dispositivo de polarização de aprendizado a partir da teoria de aprendizagem mensurável. É uma parte dinâmica da pesquisa de aprendizagem de máquinas.

É notado na figura 13 que o SVM converte espaço de entrada em um espaço de característica dimensional maior, onde os dados são separáveis.



**Figura: 13** - Representação da máquina de vetor de suporte. Fonte: Gunn (1998).

O SVM torna-se popular sempre que utiliza mapas de pixel como dados de entrada, dá exatidão equivalente a redes neurais avançadas com destaques expostos em uma atribuição de reconhecimento de caligrafia, por exemplo. Também é adotado para algumas aplicações, ou seja, exame de escrita manual, análise de rosto e assim por diante, particularmente para aplicações de classificação padrão e regressão. A Vapnik (1997) estabeleceu as máquinas de suporte de vetores e obteve proeminência devido a inúmeras características encorajadoras, como melhor desempenho observacional. As máquinas de suporte de vetores foram criadas para cuidar do problema de classificação, no entanto, nos últimos anos, eles também foram esticados para cuidar de problemas de regressão (VAPNIK; GOLOWICH; SMOLA, 1997).

De acordo com Wu et al. (2015), o pensamento essencial de Support Vector Machines é converter pontos de dados do espaço de dados para um espaço de recursos dimensional elevado, utilizando uma função de kernel para que os pontos de dados no espaço de recursos se tornem linearmente divisíveis.

1. **DESCRITORES DE TEXTURA – TD**

Esta seção especifica a definição da textura da imagem, que indica características e padrões de intensidade da imagem, sua diferenciação e descrição da textura. Em seguida, descreve possíveis abordagens para descritores de textura, como o estrutural e o estatístico, destacando o segundo, como foi usado nesta tese.

A textura é uma propriedade interna de quase todas as superfícies naturais, como a madeira, o tecido de um tecido, os padrões de areia, folhas, etc. Contém informações sobre o arranjo estrutural das superfícies e sua relação com o meio ambiente (WEI; HONG-YING, 2016). A textura representa detalhes em uma imagem, como visto na Figura 14.



**Figura 14** - Padrões de textura repetidos. Fonte: Yao et al. (2012).

A textura é um sinal efetivo para descrever estruturas de cena através de imagens que mostram um alto nível de semelhança em seu arranjo, ou seja, padrões repetidos; permitir numerosas atribuições de visão por computador, por exemplo, conhecimentos de imagem e classificação; agrupamento de cena; e recuperação visual (QUAN; XU; SUN, 2014).

De acordo com Haralick e Shapiro (1991), a principal preocupação da textura é respeitada ao arranjo espacial das intensidades de imagem e características tonais discretas. Sempre que uma área limitada da imagem tem pouca variedade de características tonais discretas, a propriedade esmagadora dessa região é o estresse cinza. E sempre que uma área limitada da imagem possui uma grande variedade de características tonais discretas, a propriedade esmagadora dessa região é textura.

# Três aspectos são vitais nesta diferenciação: a extensão das áreas limitadas; as proporções relativas das características tonais discretas; e a quantidade de características tonais discretas reconhecíveis. A textura pode ser explicada em intervalos de densidade, regularidade, grosseira, consistência de rugosidade, direcionalidade e amplitude



**Figura 15** - Amostras de textura. Fonte: Balas (2006).

Existem algumas texturas que contêm uma periodicidade sólida, chamada pseudoperiodic, uma vez que possuem disposição ou distribuição de linhas (por exemplo, duas primeiras imagens na linha superior, da esquerda para a direita na figura 15). Outros são compostos por elementos estruturais repetidos, mas carecem de periodicidade sólida ou estrutura abrangente, por exemplo, a última imagem na linha superior, da esquerda para a direita (Fig. 15). Finalmente, alguns deles têm repetição assimétrica e podem produzir efeitos de iluminação sugerindo profundidade, como a primeira imagem na linha inferior, da esquerda para a direita da figura 15 (BALAS, 2006). Todos esses tipos de texturas, precisamente pseudoperíodicas, estruturadas e assimétricas, são observados com facilidade na figura 15.

A Figura 16 destaca a última imagem da figura 15, linha inferior, da esquerda para a direita, mostrando que partes da mesma imagem podem ter uma classificação diferente, de acordo com sua textura.



**Figura 16** - Texturas de diferentes partes da mesma imagem. Fonte: Balas (2006).

A textura é universal em inúmeras imagens naturais e desempenha um papel vital como um sinal visual padrão para uma grande variedade de aplicações de exame de imagem. A classificação e a representação da textura são atribuições essenciais na imagem computacional, bem como para o reconhecimento de padrões. Foi amplamente utilizado como parte de inúmeras aplicações, incluindo investigação de imagem terapêutica, inspeção visual automatizada, reconhecimento de objetos, detecção remota e recuperação de imagem, entre outros (HADIZADEH, 2015).

No entanto, a textura é fácil de ser identificada pelos olhos humanos, é difícil ser definido em termos matemáticos. Haralick e seus colegas (1973) definem especificamente a textura de uma maneira mais rigorosa, considerando-a como um conjunto de recursos extraídos do domínio espacial para uma dada distribuição de probabilidade de escala de cinza em uma imagem.

Classificar texturas é uma questão problemática no reconhecimento de padrões e visão por computador. O compromisso é relacionar um rótulo de classe com sua correspondente divisão de textura. Nos últimos anos, foram propostas diversas abordagens de descrição de textura. Essas abordagens podem ser separadas em duas classes, ou seja, representações escassas e densas. O primeiro funciona ao reconhecer os aspectos das características, quer considerando pontos de interesse ou pontos de inspeção densos (OJALA; PIETIKAINEN; MAENPAA, 2002).

Em seguida, a descrição da característica é realizada nestes pontos de inspeção. O segundo sistema, representações densas, inclui separar elementos próximos para cada pixel em uma imagem (KHANA et al., 2015).

Outras duas abordagens possíveis para a descrição da textura são o estrutural e o estatístico. Em ambos os casos, alguns requisitos devem ser considerados, nomeadamente a invariância à posição, escala e rotação. O principal exemplo de estrutura é a transformada de Fourier da imagem. De acordo com Haralick, Shanmugam e Dinstein (1973), a abordagem estatística mais usual é a matriz de concorrência, graças ao seu melhor desempenho.

* 1. **LBP - Local Binary Patterns**

O LBP (Local Binary Patterns) é um operador para texturas considerado simples e eficiente. Ele se baseia numa ideia trivial: atribui um rótulo (representado por um número binário) a cada pixel analisado. Para atribuir esse rótulo, ele verifica se cada vizinho, do pixel em análise, em um raio considerado, extrapola, ou não, um limiar (threshold) que é determinado pelo valor do pixel central, cujo rótulo está sendo determinado. Cada vizinho que extrapolar esse limiar (threshold), ou se igualar a ele, será marcado com o dígito binário 1, caso contrário, será marcado com o dígito binário 0. A junção desses dígitos binários representará o pixel central, lembrando que é importante seguir uma lógica para determinar qual o dígito mais significativo, ou, em outras palavras, à partir de que vizinho o número começará a ser formado, já que a vizinhança é circular. Também é preciso observar que quando se determina o dígito mais significativo, automaticamente está sendo atribuído um peso a cada um dos dígitos. A soma da multiplicação de cada dígito pelo seu respectivo peso retornará o que é referenciado como código LBP (PIETIKÄINEN et al., 2011, p. 4-5).

Ainda segundo Pietikäinen et al. (2011, p. 4), possivelmente, a propriedade mais importante do operador LBP, em aplicações do mundo real, é a sua invariância, percebida quando acontecem variações monotônicas nos níveis de cinza, que podem ser causadas por variações na iluminação do ambiente, no momento da captura da imagem. Outra característica igualmente importante é a sua simplicidade computacional, que torna possível enfrentar desafios nos quais as imagens precisam ser analisadas em tempo real. A Figura 17 mostra uma aplicação do LBP.

 **Figura 17**: Aplicação do LBP (Fonte: Adaptado de (PIETIKÄINEN et al., 2011, p. 14))

1. **MATERIAIS E MÉTODOS**

Com o fim de colaborar na elaboração do sistema de identificação anatômica de madeira nativa por meio de comparação de imagens, que está sendo desenvolvido pela FATEC de São José do Rio Preto em parceria com a Polícia Ambiental, foram geradas imagens de amostras de madeira nativa, sendo trinta imagens de cedrinho *“Erisma uncinatum”*, trinta imagens de cedro rosa “*Cedrela fissilis”*, trinta imagens de garapa *“Apuleia leiocarpa”* e trinta imagens de cupiúba *“Goupia glabra”*, perfazendo cento e vinte imagens de madeira oriundas da Floresta Amazônica. As imagens foram geradas pela Polícia Ambiental a partir de amostras das aludidas espécies de madeira nativa com uso do equipamento denominado Dino-Lite, modelo AM-313T, com aumento aproximado de dez vezes. Para a obtenção das imagens, a metodologia utilizada consiste em retirar uma amostra da madeira a ser analisada e, realizando-se um corte perpendicular ao comprimento da árvore com lâmina bem afiada, retira-se a imagem com o equipamento acima indicado na superfície seccionada com maior polimento.

O resultado esperado é uma imagem nítida, que permita a observação de características anatômicas próprias, tais como a presença e forma dos parênquimas, a disposição, quantidade e características dos poros, a existência e características de raios, disposição das fibras, bem como coloração e tonalidade da madeira, informações estas que possibilitam classificar a imagem de acordo com sua respectiva espécie de madeira em percentual de semelhança cientificamente aceito.

****

****

**Figura 17** - Dino-Lite modelo AM-313T

Após a aquisição da imagem, o processo de aprimoramento iniciado com a ajuda do software Weka. Nesta fase, as imagens precisavam ser uniformes e um processo de padronização ocorreu através do referido sistema operacional.

O sistema de processamento de imagem funciona da seguinte forma. Os recursos são extraídos de imagens em cores, tratando cada canal de imagem a cores (Vermelho-Amarelo-Verde) como uma imagem monocromática e transformando suas informações de forma em pixels de superfícies.

|  |  |
| --- | --- |
| Aquisição de imagem | Ele é usado para obter a imagem a ser processada a partir de um determinado objeto. |
| Melhoria de imagem | É necessário elevar a qualidade da imagem do objeto para um processo posterior. |
| Segmentação | Para dividir a imagem em áreas específicas de interesse. |
| Extração de recursos | Técnica utilizada para calcular os parâmetros de taxas ou medidas que diferenciam cada objeto do outro. |
| Classificação | É apropriado para a determinação do que representa cada objeto. |

Tabela 2 - Etapas do processamento de imagem. Fonte: Pham e Alcock (2003)

Para realizar a classificação visual das espécies, as amostras de dimensão foram separadas em 4 grupos:

**A:Cedrinho**

**B:Cedro Rosa**

**C:Cupiba**

**D:Garapa**

Inicialmente foi realiza um pré-processamento das imagens, constando de um recorte, transformação em escala de cinza. Através do recorte os tamanhos das imagens foram padronizados nas dimensões 160x120 pixels, posteriormente os três canais de imagens RGB (Red, Green and Blue) foram transformados num canal único contendo as informações em escala de cinza. Para realização da extração de características foram utilizados dois descritores de textura, sendo Haralick e LBP (Local binary pattern).

Tanto o Haralick quanto o LBP conseguiram identificar a relação de um deteminado pixel com sua vizinhança. Esta etapa possui, dois objetivos, identificar a relação de um determinado pixel com seus vizinhos próximos, reduzir a dimensionalidade do problema e identificar as características mais importantes de uma imagem. Desta da forma, o vetor de características da imagem contém 12 descritores de haralick e 59 descritores do tipo LBP, e na última coluna indica a classificação da amostra (A=1, B =2, C = 3, D = 4).

Portanto, a base de dados é composta do vetor de característica de todas as imagens coletadas. A fim de realizar a classificação através de softwareWEKA, a base de dados foi escrita no formato .arff.

# Referências

Affonso, C., Sassi, R. J., & Barreiros, R. M. (2015). Biological image classification using rough-fuzzy artificial neural network. *Expert Systems with Applications, 42* (24), 9482–9488. doi: 10.1016/j.eswa.2015.07.075 .

Affonso, C. *et al.* Deep learning for biological image classification. **Expert Systems with Applications**.

ANTUNES, E. Processamento de Imagens: Uma abordagem interdisciplinar aplicada à correção de prognósticos meteorológicos. Pelotas, RS, Brasil, 1999.

BURGER, L.M.; RICHTER, H.G. **Anatomia da madeira**. São Paulo: Nobel, 1991. 154 p.

COTRIM, D. Arquitetura de indexação aplicada a servidores PACS para processamento de imagens. Dissertação (Mestrado) — Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brasil, 2007.

EFFORD, N. Digital Image Processing: A Pratical Introduction Using Java. 1. ed. Leeds, England: Addison Wensley, 2000. ISBN 0201596237.

ESQUEF, I. Técnicas de Entropia em Processamento de Imagens. Dissertação (Mestrado) — Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, RJ, Brasil, 2002.

GONZALEZ, R.; WOODS, R. Digital Image Processing. 3. ed. Reading, MA, USA:

Addison-Wesley, 1992. ISBN 0-201-50803-6.

INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS RENOVÁVEIS. **Padronização da nomenclatura comercial brasileira das madeiras tropicais amazônicas**. Brasília, 1991. 85 p.

JAIN, A. K. Fundamentals of Digital Image Processing. Englewood Cli\_s, NJ, USA: Prentice Hall, 1989. ISBN 0-13-336165-9.

MANIERI, C. **Manual de Identificação das principais Madeiras Comerciais Brasileiras.** São Paulo: Companhia de promoção de Pesquisa Científica e Tecnológica do Estado de São Paulo, 1989. 240p.

MASCARENHAS, N.; VELASCO, F. Processamento Digital de lmagens. 1. ed. Buenos Aires, Argentina: Kapelusz, 1989.

Oppenheim, A.V.; SCHAFER, R.W. **Digital Signal Processing**. London: Prentice-Hall International, 1975. x p.

SILVA, N. R. *et al.* Automated classification of wood transverse cross-section micro-imagery from 77 commercial Central-African timber species. **Annals of Forest Science**.

TURKSEN, I. Interval value fuzzy sets based on normal form. Fuzzy Sets and Systems, Amsterdam, Netherlands, v. 20, p. 191–210, 1986. ISSN 0165-0114.

VIEIRA, F. H. A.; PROCESSAMENTO DE IMAGENS ATRAVÉS DE APRENDIZADO DE MÁQUINAS PARA A CLASSIFICAÇÃO DA QUALIDADE DA MADEIRA, Tese apresentada à Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista, 2016.

**Apêndice A - KNN**

import numpy as np

# Euclidean distance

def distance (x,y):

return np.sqrt(np.power(x-y,2).sum(axis=1))

def knn (in\_train, out\_train, in\_test, k):

size\_in\_train = in\_train.shape

size\_out\_train = out\_train.shape

size\_in\_test = in\_test.shape

#The labels array that will be returned

labels = np.zeros ([size\_in\_test[0], size\_out\_train[1]])

for i in range(size\_in\_test[0]):

# Computing the distance from the sample test to the training set

rpt\_test = np.tile (in\_test[i,:], (size\_in\_train[0], 1))

dists = distance (rpt\_test,in\_train)

# Sorting the distances and getting the k nearest neighbors

index\_sort = np.argsort (dists)

pos\_labels = index\_sort[:k]

closeness = out\_train [pos\_labels]

# The final label will be the highest value in the row

labels[i] = closeness.sum(axis=0)

return labels