

PESQUISAR

Acesso livre



Abordagens de rede neural de aprendizado profundo e aprendizado de máquina para classificação e segmentação de defeitos de textura de couro multiclasse

Praveen Kumar Moganam e Denis Ashok Sathia Seelan*

Abstrato

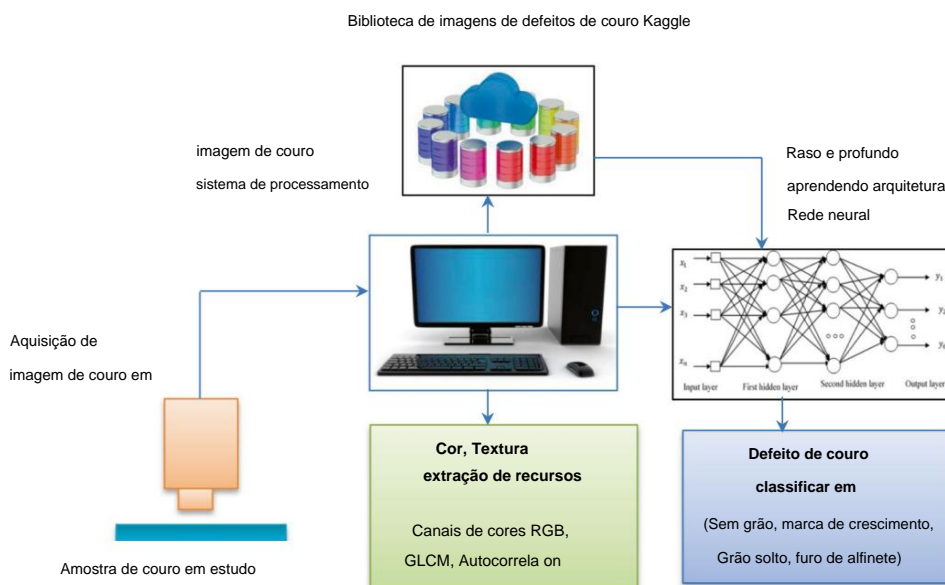
As indústrias de couro modernas estão focadas em produzir produtos de couro de alta qualidade para sustentar a competitividade do mercado. No entanto, vários defeitos de couro são introduzidos durante vários estágios do processo de fabricação, como manuseio de material, curtimento e tingimento. A inspeção manual de superfícies de couro é subjetiva e inconsistente por natureza; portanto, os sistemas de visão de máquina foram amplamente adotados para a inspeção automatizada de defeitos de couro. É necessário desenvolver algoritmos de processamento de imagem adequados para localizar defeitos de couro, como marcas de dobra, marcas de crescimento, granulação, granulação solta e pinhole devido ao padrão de textura ambígua e natureza minúscula nas regiões localizadas do couro. Este artigo apresenta uma abordagem baseada em rede neural de aprendizado profundo para localização e classificação automática de defeitos de couro usando um sistema de visão de máquina. Neste trabalho, redes neurais convolucionais populares são treinadas usando imagens de couro de diferentes defeitos de couro e uma técnica de mapeamento de ativação de classe é seguida para localizar a região de interesse para a classe de defeito de couro. Redes neurais de convolução, como Google net, Squeeze-net, ResNet, fornecem melhor precisão de classificação em comparação com as arquiteturas de rede neural de última geração e os resultados são apresentados.

Palavras-chave: Redes neurais de convolução, Classificador de aprendizado de máquina, Defeitos de couro, Classificação multiclasse, Mapa de ativação de classe, Segmentação

*Correspondência: denisashok@vit.ac.in

Departamento de Automação de Projetos, Laboratório de Sistemas Ciberfísicos, Escola de Engenharia Mecânica, Vellore Institute of Technology, Vellore 632014, Índia

Resumo gráfico



1. Introdução

Os fabricantes e designers modernos de couro têm dado grande atenção à percepção estética, aparência visual e toque das roupas de couro, pois isso afeta a decisão de compra. No entanto, vários defeitos são introduzidos na superfície do couro durante os processos de pré-curtimento e pós-curtimento nas indústrias de couro. Assim, a identificação e classificação dos defeitos do couro é um processo essencial para a manutenção da qualidade dos produtos acabados. Como os métodos de inspeção manual são lentos, propensos a erros e trabalhosos, técnicas de inspeção automatizadas baseadas em visão de máquina são amplamente adotadas para melhorar a produtividade do processo de inspeção de couro [1]. Devido ao padrão de textura ambíguo e tamanho minúsculo do defeito, é difícil distinguir o defeito localizado e o fundo nas imagens de couro. Portanto, existe a necessidade de desenvolver uma abordagem de processamento de imagem adequada para melhorar a classificação e percepção dos defeitos do couro.

Várias técnicas de processamento de imagem foram propostas por muitos pesquisadores para classificação de couro, identificação de defeitos e classificação. A inspeção de qualidade para classificação é uma etapa importante na avaliação da área utilizável de couros. Cada peça de couro é classificada com base em seu valor de corte efetivo, que é decidido levando em consideração o número, tamanho e localização dos defeitos superficiais [2]. Técnicas de processamento de imagem em escala de cinza usando limiarização e operações morfológicas são aplicadas para defeitos

aplicações de detecção [3]. Um método de identificação baseado em histograma é proposto para detectar imagens de couro defeituosas [4]. A detecção de bordas juntamente com operações morfológicas é aplicada às imagens de couro para segmentar os locais de defeitos nas imagens de couro [5]. A técnica de análise de textura usando transformada wavelet fornece uma análise espacial coletiva de regiões de pixel locais para detecção de defeitos de couro [6]. Um algoritmo de limiarização multinível com o método de extração de características de textura é proposto para segmentar regiões defeituosas e não defeituosas de couro para quantificar objetivamente os defeitos da superfície do couro [7]. Sobral e cols. introduziu um novo método baseado em wavelets usando bancos de filtros otimizados para detecção de defeitos em couro [8]. Uma abordagem de otimização com um processo de filtragem é aplicada para isolar as regiões defeituosas do fundo complexo e não homogêneo, analisando sua estrutura fortemente orientada [9]. Para o processo de detecção e classificação de defeitos, vários algoritmos de processamento de imagem são empregados para fornecer as descrições quantitativas de imagens de couro defeituosas e não defeituosas, vários descritores como estatísticas de primeira ordem, características de contraste, descritores de Haralick, transformada de Fourier e Cosseno, momentos de Hu com informações sobre a intensidade, padrões binários locais, características de Gabor são extraídas para localizar as posições do defeito na superfície do couro [10]. Os recursos de Haralick são derivados da matriz de co-ocorrência de nível de cinza (GLCM), que extrai os padrões locais na imagem e conta sua distribuição



Fig. 1 Sistema de inspeção de visão de máquina para identificação de defeitos em amostras de couro

Tabela 1 Principais especificações da câmera usada para aquisição de imagens de couro

Especificações	Descrição
Faço	Basler acA 4600
Tipo de sensor	CMOS
Tamanho do sensor	6,5 mm × 4,6 mm
Resolução (A × V)	4608 pixels × 3288 pixels
Resolução	14 MP
Tamanho do pixel (A × V)	1,4 μm × 1,4 μm
Taxa de quadros	10 fps
Mono/cor	cor
Interface	USB 3.0

em toda a imagem. Eles fornecem uma boa codificação discriminativa da informação textural e baseada em gradiente, na forma de valores de características [11]. Um método de análise de textura usando recursos estatísticos wavelet e recursos de matriz de co-ocorrências wavelet, como entropia, energia, contraste, correlação, proeminência do cluster, desvio padrão, média e homogeneidade local é proposto para a classificação de defeitos de couro [12]. Modelos baseados em cores e análise de textura baseada em matriz de coocorrência são relatados para detecção de defeitos em couro cru [13]. Embora as abordagens de processamento de imagem digital sejam aplicadas para aplicações de inspeção de defeitos de couro, a classificação de precisão é limitada devido à presença de ruído.

Recentemente, os métodos de aprendizado de máquina e aprendizado profundo ganharam atenção para aplicações de classificação, detecção e segmentação de imagens. Kwak et al. propuseram uma árvore de decisão sequencial de três estágios para a classificação de defeitos como linhas, furos, manchas, desgastes e nós [14]. Viana e cols. apresentou uma avaliação empírica da máquina de vetores de suporte contra AdaBoost e MLP, para resolver o problema de classificação de defeitos de couro [15]. Classificação supervisionada usando o multi-layer perceptron (MLP), árvores de decisão (DT), SVM, Naïve Bayes, KNN,

e classificadores Random Forest (RF) foram usados para classificar as regiões de couro defeituosas e não defeituosas [16]. O classificador de rede neural é proposto por redes neurais per ceptron multicamadas para reconhecer defeitos de couro como corte aberto, corte fechado e Fly Bite [17]. Amorim e cols. apresentou técnicas de análise linear discriminada para redução de atributos para quatro classificadores diferentes, como C4.5, KNN, Naïve Bayes e classificadores SVM para classificação de defeitos de couro [18]. Com os avanços recentes em computação e unidades de processamento gráfico, redes neurais de aprendizado profundo são desenvolvidas para aplicações de inspeção automatizadas [19]. As arquiteturas ResNet e VGG baseadas em redes neurais de convolução são capazes de aplicações automatizadas de inspeção de superfície e classificação de imagem usando aprendizado de transferência [20]. Liong et al. propuseram um sistema integrado de visão de máquina usando uma rede neural artificial e uma rede neural de aprendizagem profunda para classificação de defeitos de couro [21]. A abordagem de aprendizado profundo baseada em rede neural de convolução de região é usada para detecção de defeitos e segmentação de regiões defeituosas na imagem de couro [22]. Com base na percepção visual-tátil dos consumidores, a rede neural de retropropagação é desenvolvida para selecionar os materiais de couro adequados para fabricar os produtos de couro especificados pelo usuário [23].

Constata-se que muitos trabalhos de pesquisa têm contribuído para a detecção e classificação de defeitos em couro. As abordagens de classificação existentes limitaram a percepção humana usando imagens de couro e a precisão da classificação também é limitada devido à imprecisão, aleatoriedade e tamanho dos defeitos de couro no padrão de textura de fundo da superfície de couro. A fim de fornecer maior precisão na classificação de defeitos do couro e na percepção humana; este artigo apresenta redes neurais de convolução de aprendizado profundo e abordagens de classificação de aprendizado de máquina para classificação multiclasse e segmentação de defeitos de couro. O desempenho de classificação de classificadores de aprendizado profundo e aprendizado de máquina de última geração é comparado para os conjuntos de dados de couro com defeitos de textura e os resultados são apresentados neste artigo.

2 Sistema de inspeção de couro baseado em visão de máquina

Uma superfície típica de couro consiste em diferentes tipos de defeitos, como cicatrizes, marcas de crescimento, grão, grão solto, furos e marcas de dobra. Um sistema de visão de máquina composto por uma câmera de alta resolução (BASLER acA4600), sistema de iluminação, sistema de computação com um software de processamento de imagem (MATLAB Versão. 2020a) é estabelecido no presente trabalho para identificar e classificar os defeitos do couro e é mostrado na Fig. 1.

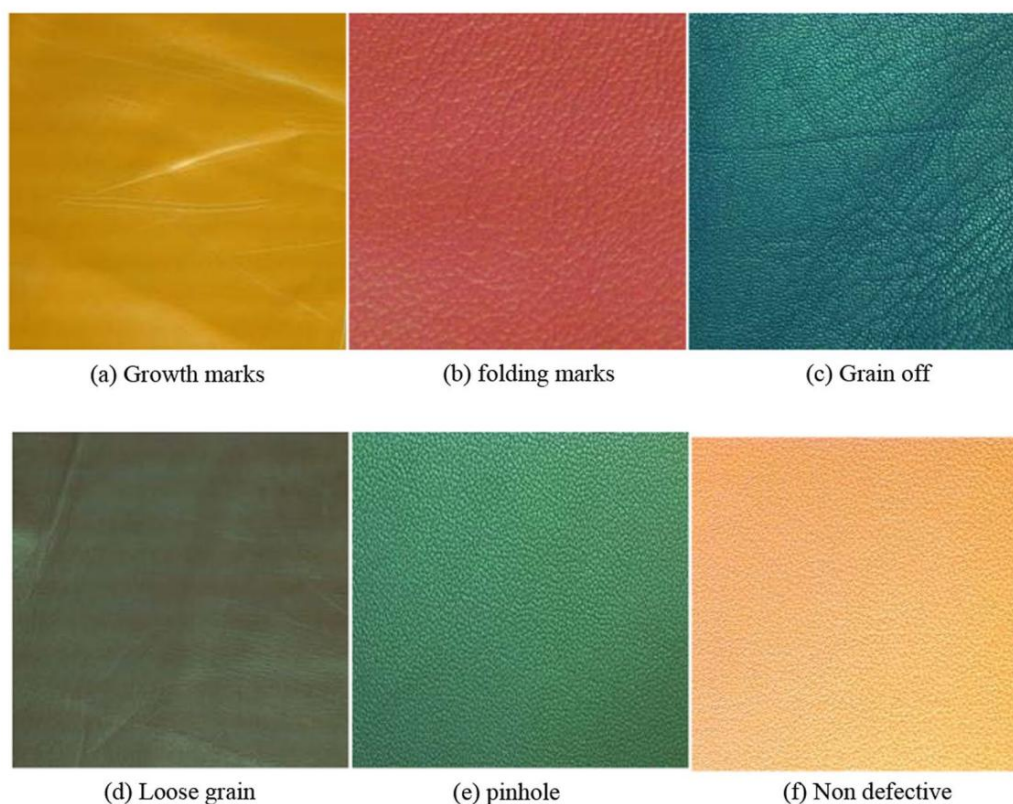


Fig. 2 Amostras de superfícies de couro com cores e defeitos diferentes

Uma câmera de alta resolução com resolução de 14 MP é utilizada para adquirir as imagens da superfície do couro com resolução de 4608x3288 pixels. A Tabela 1 mostra as especificações da câmera utilizada no sistema de visão de máquina.

A iluminação desempenha um papel importante nas aplicações de inspeção visual para iluminar o objeto de interesse. Neste trabalho, um sistema de iluminação por fibra ótica é utilizado para fornecer iluminação uniforme na superfície do couro. A magnitude da luminância é medida usando um medidor Lux e é controlada usando um botão controlador de luz.

2.1 Aquisição de imagem de couro

Um conjunto de dados abrangente de 3600 imagens de couro é desenvolvido com diferentes defeitos, como marcas de dobra, granulação, orifício, marcas de crescimento, granulação solta e superfícies de couro sem defeitos. Ele é colocado no ambiente de ciência de dados aberto Kaggle para explorar o aprendizado de máquina adequado e a técnica de processamento de imagem baseada em aprendizado profundo para classificar os defeitos do couro. A Figura 2 mostra as amostras de imagens de couro com diferentes defeitos de couro.

Pode-se notar na Fig. 2a, b que as marcas de dobras e marcas de crescimento têm uma melhor percepção visual da mudança

em cor e textura em comparação com outros defeitos.

Além disso, descobriu-se que o grão de, pinhole e grão solto têm um padrão de textura mais fino, conforme identificado na Fig. 2c-e, respectivamente.

2.2 Defeitos na textura do couro

As imagens de couro de diferentes cores e defeitos como mostrado na Fig. 2 são processadas para obter os mapas de intensidade de escala de cinza para análise das variações de textura devido a diferentes defeitos de couro e os resultados são mostrados na Fig.

3. Marcas de dobra, grão de, pinhole e grãos soltos têm uma textura grosseira e perturbam o padrão de textura regular, o que leva a muitas variações abruptas e picos na intensidade dos pixels, conforme mostrado na Fig. 3a-d, respectivamente.

A marca de crescimento e o couro não defeituoso mostraram textura mais fina e variação de intensidade uniforme é vista na Fig. 3e, f respectivamente. A percepção visual dos defeitos do couro é limitada pelo padrão de textura ambígua e pela natureza minúscula dos diferentes defeitos do couro. A fim de distinguir os tipos de defeitos de couro, há a necessidade de desenvolver algoritmos de processamento de imagem adequados para classificação de diferentes defeitos de couro.

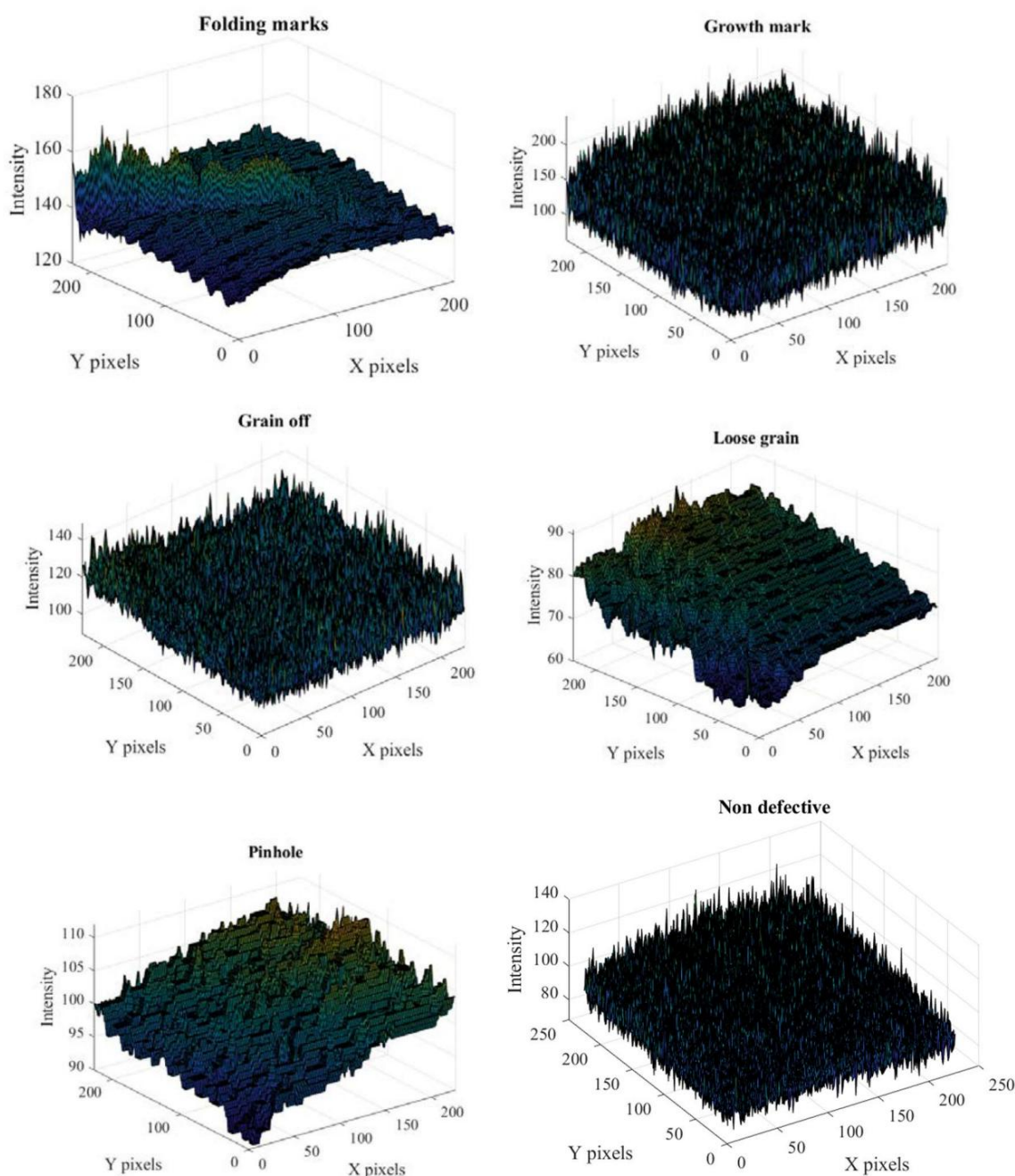


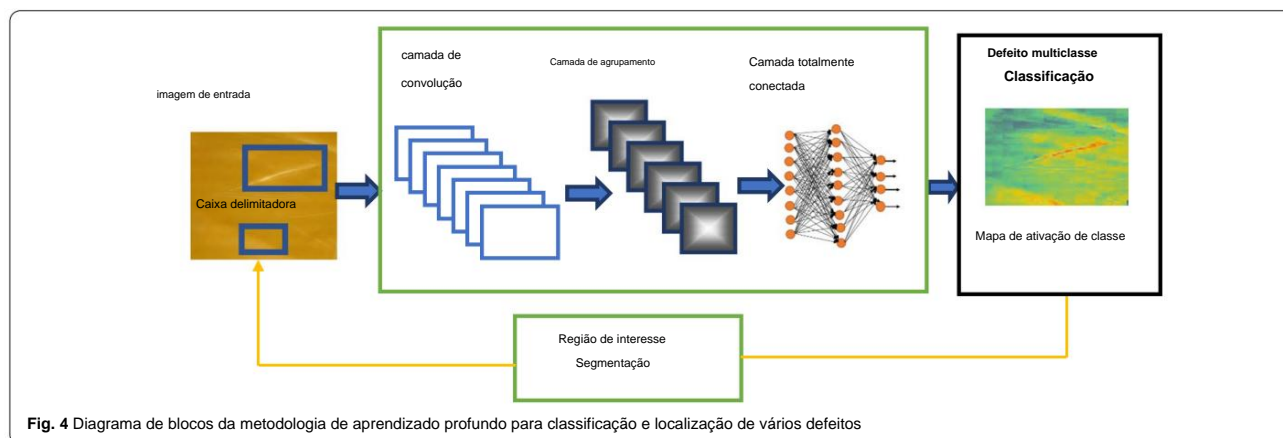
Fig. 3 Mapa de intensidade de cinza de diferentes imagens de couro com diferentes defeitos de couro

3 Abordagem de rede neural de aprendizagem profunda para classificação e localização de defeitos de couro

A fim de reduzir o erro na detecção e classificação multiclasse de defeitos de textura de couro, este trabalho apresenta uma abordagem de rede neural convolucional de aprendizado profundo usando arquiteturas de rede neural convolucional de última geração como Alexnet, VGG-16, Google

net, Squeeze Net, ResNet-50. A Figura 4 mostra a estrutura da abordagem proposta para classificar e rotular imagens de couro como couro não defeituoso, grão solto, grão de, marcas de crescimento, pinhole e marcas de dobra.

Usando os modelos de rede neural de aprendizado profundo desenvolvidos, o mapa de ativação de classe é gerado para identificar a região de interesse da classe de imagem de couro.



Os detalhes da arquitetura de rede, as camadas são explicadas nas subseções a seguir.

3.1 Preparação e pré-processamento do conjunto de dados da imagem de couro

Um conjunto de dados de 3600 imagens de couro consistindo em couro não defeituoso, grão solto, grão de, marcas de crescimento, furos e marcas de dobramento é usado para treinar redes neurais. Todas as imagens de couro são pré-processadas usando equalização de histograma para lidar com as variações de iluminação durante a aquisição da imagem e redimensionadas para 227×227. Para avaliar o desempenho da segmentação de regiões defeituosas na imagem de couro, as imagens de couro real com regiões defeituosas rotuladas à mão são mantidas no conjunto de dados. Uma abordagem de validação cruzada quintupla é seguida no presente trabalho, na qual o conjunto de dados de 3600 amostras de imagens é dividido em 5 dobras mutuamente exclusivas e exaustivas de 720 imagens de couro. Repetidamente, uma dobra é selecionada e designada como conjunto de teste, todas as outras imagens de couro restantes (80% dos dados) são consideradas como conjunto de treinamento.

3.2 Rede neural convolucional de aprendizagem profunda arquiteturas

No presente Padrão, arquiteturas de redes neurais convolucionais pré-treinadas, como Alexnet, VGG-16, ResNet, Google net, Inception Net e Squeeze net, são consideradas para a aplicação de classificação de defeitos de couro multiclasses, pois são relativamente estabelecidas e provaram sua capacidade para objetos aplicações de detecção e classificação multiclasses. A Figura 5 mostra as arquiteturas de redes neurais convolucionais de aprendizado profundo consideradas no presente trabalho. Um modelo típico de rede neural convolucional contém camada de convolução, camada de pooling e camada totalmente conectada. Pode ser visto que a arquitetura de trabalho de rede de complexidade aumenta com concatenação, canais paralelos e realimentação, conforme mostrado na Fig. 5c, d para rede inicial e Resnet, respectivamente. As camadas de convolução

estão associados a diferentes parâmetros, como pesos, tamanho do kernel, passo, preenchimento etc. Mais detalhes sobre arquiteturas de aprendizado profundo podem ser estudados na literatura [24].

3.2.1 Camada de convolução

A camada de convolução fornece extração automatizada de recursos de determinadas imagens com localizações espaciais específicas usando filtros de número de tamanhos diferentes. Um mapa de recursos não sobreposto é obtido como uma saída usando a operação de convolução entre os pesos do filtro e a saída da camada convolucional anterior, conforme dado pela Eq. (1).

$$I_x = f(\sum_{j=1}^J \sum_{l=1}^L w_{l,j} x_{l,j} + b_l) \quad (1)$$

onde (J, L) denota o tamanho dos filtros, J é a altura dos filtros e L é a largura dos filtros. b_l denota o viés da camada convolucional. $x_{l,j}$ denota a saída da camada convolucional anterior. $w_{l,j}$ denota o peso da camada convolucional. $f(\cdot)$ é a função de ativação não linear e a função de ativação ReLU é selecionada e é mostrada como Eq. (2).

O tamanho do mapa de recursos depende de vários parâmetros, incluindo o tamanho de entrada, tamanho do filtro, profundidade da pilha do mapa, preenchimento de zero e passo.

$$M_x = \frac{I_x - K_x}{S_x}, \quad m_{y_l} = \frac{I_y - K_y}{S_y}, \quad (2)$$

onde (M_x, M_y), (I_x, I_y), (K_x, K_y) indicam o tamanho do mapa, tamanho da entrada, tamanho do kernel, respectivamente, e (S_x, S_y) indicam o passo na linha e na coluna. O número de mapas de recursos depende do número de filtros e de seu tamanho. Em uma rede neural de convolução de aprendizagem profunda típica, o número de recursos com o aumento no número de camadas de convolução e os filtros associados.

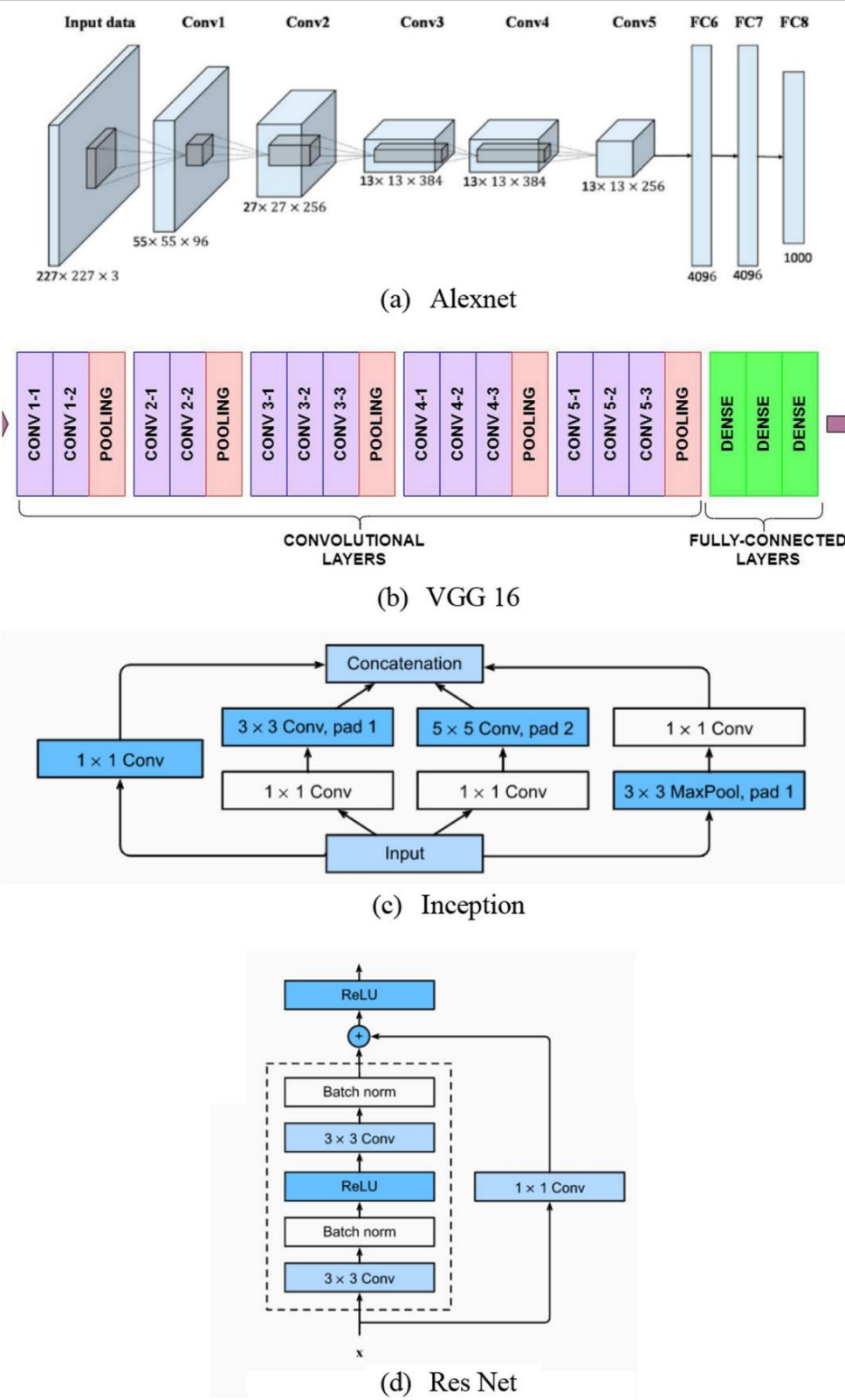
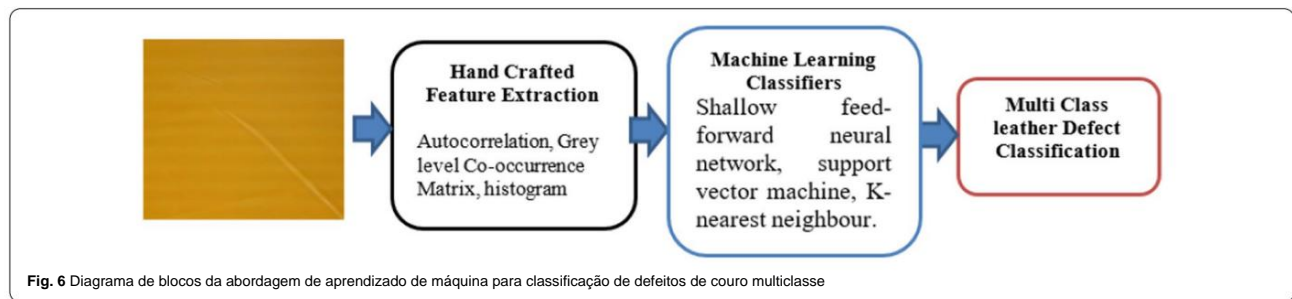


Fig. 5 Diagrama de blocos de redes neurais convolucionais de aprendizado profundo padrão



3.2.2 Função de ativação

Para aprender a aproximação universal de valores de entrada e classes de saída como função contínua em um espaço euclidiano, uma função de ativação adequada é essencial. Como as funções sigmoide e tangente hiperbólica sofrem o problema do gradiente de fuga, uma unidade retilínea é usada como função de ativação e retorna o mesmo valor (x) dos mapas de recursos fornecidos como saída (x) se sua magnitude for maior que zero.

3.2.3 Camada de agrupamento

A fim de evitar o overfitting e reduzir a dimensionalidade dos mapas de recursos, a camada de pooling executa uma amostragem reduzida do mapa de recursos de entrada usando uma função de janela $u(n,n)$. As estratégias de pooling máximo e pooling médio geralmente são seguidas na camada de pooling.

$$x_m = \max_{N \times N} (xlu(n, n)) \quad (3)$$

onde $u(n,n)$ é a função janela, que é aplicada para calcular o valor máximo de x_{ij} na vizinhança.

3.2.4 Camada totalmente conectada

A camada totalmente conectada na rede neural de aprendizado profundo recebe um vetor de recursos da camada de agrupamento máximo anterior e é treinada para a classificação multiclasse de determinada imagem de couro usando os pesos associados e uma função de ativação reduzindo uma função de perda. Neste trabalho, uma camada totalmente conectada com seis números de neurônios de saída é configurada para fornecer a saída categórica, como couro não defeituoso, grão solto, grão de, marcas de crescimento, pinhole e marcas de dobra usando uma técnica de codificação.

Mais detalhes sobre o treinamento da rede neural para a classificação multiclasse são descritos na Seção 4.2.

A saída típica de um neurônio em uma camada totalmente conectada para o vetor de características da camada de agrupamento máximo $xm\tilde{y}1$ é dada pela Eq. (4).

$$v^m = f(wm xm\tilde{y}1 + bm) \quad (4)$$

onde bm denota o viés da camada totalmente conectada. wm denota os pesos da camada totalmente conectada. $xm\tilde{y}1$ denota a saída da camada max-pooling anterior. $f(.)$ é a função de ativação.

3.3 Visualização da região de interesse do defeito localização

Neste trabalho, o mapeamento de ativação de classe ponderada por gradiente é seguido, que é a soma ponderada de cada canal do mapa de recursos para identificar as regiões discriminativas específicas da imagem de couro fornecida couro não defeituoso, grão solto, grão de, marcas de crescimento, Pinhole e marcas de dobramento. Aqui, o mapa de ativação de classe com valores de pontuação para a classe do couro 'c' na localização espacial (x,y) da imagem é gerado usando o k -ésimo canal do mapa de recursos e o peso correspondente wck conforme dado por

$$Mc(i, j) = \sum_{k=1}^K wck \cdot fk(x, y) \quad (5)$$

onde Mc é o mapa de ativação da classe c e wck representa o k -ésimo peso da camada totalmente conectada da classe c . Como a parte da imagem com maior pontuação influencia a classe correspondente, uma abordagem de limiarização é seguida para a seleção da região de interesse na imagem original.

$$T(x, y) = \begin{cases} 0 & \text{se } M(x, y) > \text{limite} \\ 1 & \text{se } M(x, y) < \text{limite} \end{cases} \quad (6)$$

Posteriormente, a região de interesse $ROI(x,y)$ é obtida com limite $T(x,y)$ para indicar a região discriminativa na imagem $I(x,y)$:

$$ROI(x, y) = I(x, y) \cdot T(x, y) \quad (7)$$

Uma caixa delimitadora é gerada a partir da região de interesse da imagem dada e é comparada com t caixas delimitadoras de base nas imagens de couro de base.

Tabela 2 Descrição dos recursos de textura GLCM usados para identificação de defeitos de couro [10]

Parâmetros de textura GLCM	Descrição
Entropia = $\sum_{i=0}^{ng-1} \sum_{j=0}^{ng-1} \hat{p}_{ij} \hat{y} \log \hat{p}_{ij}$	Medida da aleatoriedade estatística da superfície do couro
Correlação = $\frac{\sum_{i=0}^{ng-1} \sum_{j=0}^{ng-1} \hat{p}(i,j) \hat{y}_i \hat{y}_j}{\hat{y}_x \hat{y}_y}$	Medida de dependência linear de níveis de cinza de pixels vizinhos
Contraste = $\sum_{i,j=0}^{ng-1} \hat{p}_{i,j} (i - j)^2$	Medida da intensidade entre um pixel e seu vizinho
Energia = $\sum_{i,j=0}^{ng-1} \hat{p}_{i,j}^2$	Medida de ordenação de pixels
Homogeneidade = $\sum_{i,j=0}^{ng-1} \frac{\hat{p}_{i,j}}{1 + (\hat{y}_i - \hat{y}_j)^2}$	Medida de suavidade da distribuição de nível de cinza
Dissimilaridade = $\sum_{i,j=0}^{ng-1} \hat{p}_{i,j} i - j $	Medida da distância entre pares de pixels
Meanµi = $\sum_{i,j=0}^{ng-1} i(\hat{p}_{i,j})$	Medida da intensidade média de todos os pixels
Variância $\hat{y}_2 = \sum_{i,j=0}^{ng-1} \hat{p}_{i,j} (i - \mu_i)^2$	Medida de dispersão de distribuição de nível de cinza de pixels

4 Abordagens baseadas em aprendizado de máquina para classificação de defeitos multiclasse de defeitos de couro

A classificação e localização de defeitos de couro multiclasse baseada em redes neurais de aprendizado profundo propostas são comparadas com as abordagens de aprendizado de máquina como rede neural feed-forward rasa (SFFNN), máquina de vetores de suporte (SVM), K-vizinho mais próximo (KNN). Essas abordagens de aprendizado de máquina requerem técnicas manuais de extração de recursos das imagens de couro para a classificação de defeitos de couro. As etapas típicas seguidas na classificação de defeitos de várias classes com base em aprendizado de máquina de defeitos de couro são mostradas na Fig. 6.

Os detalhes da extração de recursos e aprendizado de máquina classificadores são fornecidos nas seções subsequentes.

4.1 Extração de características artesanais de imagens de couro Neste trabalho, as características de cor e textura são extraídas de 3600 imagens de couro para cada classe, como couro não defeituoso, grão solto, grão de, marcas de crescimento, pinhole, marcas de dobra usando histograma de cores, relação Autocor e GLCM.

4.1.1 Histograma de cores Como o couro tem cores diferentes e tem variações de cor nos locais defeituosos da imagem do couro, ele é analisado usando histograma no espaço de cores RGB para entender as variações de intensidade nos canais de vermelho, verde e azul conforme dados pela Eq. (8)

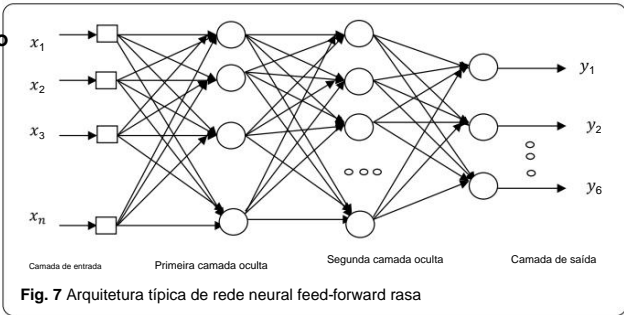


Fig. 7 Arquitetura típica de rede neural feed-forward rasa

f(i, j) = {fR(x, y), fG(x, y), fB(x, y)} (8)

onde i=1,2, ..., M/m e j=1,2, ..., N/n Além disso, um histograma colorido é aplicado para o bloco de imagem dado e valor de intensidade dominante do canal vermelho, verde e azul (Rmax, Gmax, Bmax) é extraído como o recurso de cor da imagem do couro, conforme indicado abaixo:

C máx(i, j) = {máx fR(x, y), máx fG(x, y), máx fB(x, y)} (9)

Aqui , Cmax refere-se à magnitude das contagens dos canais vermelho, azul e azul da imagem de couro. Esses recursos de cor ajudarão a quantificar as mudanças de cor da imagem do couro com os defeitos do couro.

4.1.2 Autocorrelação

Como a magnitude da função de autocorrelação é útil para descrever a perturbação na textura regular

padrão devido à presença de defeitos de couro na superfície do couro, a função de autocorrelação é calculada para a imagem do couro para medir sua aspereza devido a defeitos de couro. Aqui, a imagem de couro colorida $f(x,y)$ é convertida em imagem em escala de cinza e uma função de autocorrelação bidimensional da imagem de couro dada $fg(x,y)$ é calculado usando a seguinte Eq. (3).

$$G(a, b) = \sum_x \sum_y \sum_{x+a} \sum_{y+b} fg(x, y) \cdot fg(x+a, y+b) \quad (10)$$

onde $G(a,b)$ é a função de autocorrelação para a escala de cinza e a e b representam o atraso típico da posição x e y correspondente.

4.1.3 Matriz de coocorrência de nível de cinza

A matriz de co-ocorrência de nível de cinza fornece informações importantes para a compreensão da variação no padrão de textura devido ao tipo de defeitos de couro em marcas de dobra, grão de, pinhole, marcas de crescimento e grão solto. A Matriz de Coocorrência de Nível de Cinza (GLCM) para a imagem de couro fornecida é construída contando todos os pares de uma referência e pixel vizinho separados por um deslocamento (d) tendo os níveis de cinza i e j na orientação relativa especificada (\tilde{y}) conforme abaixo:

$$p[i, j|d, \tilde{y}] = n_{ij} \quad (11)$$

onde n_{ij} é o número de ocorrências de referência e pixels vizinhos (i,j) situados no deslocamento (d) na imagem. $p[i,j]$ é a matriz de co-ocorrência de nível de cinza e é calculada para a imagem em escala de cinza em quatro orientações diferentes ($\tilde{y}=0^\circ, 45^\circ, 90^\circ$ e 135°) e deslocamentos ($d=\tilde{y}3, \tilde{y}2, -1, 0, 1, 2, 3$). O número de linhas e colunas da matriz de coocorrência $p[i,j]$ é igual ao número de níveis de cinza distintos (n). Para reduzir a carga computacional de calcular GLCM para a imagem fornecida, o nível de cinza foi definido a 64. Além disso, os elementos de $P[i,j]$ são normalizados dividindo cada entrada pelo número total de pares de pixels.

A Tabela 2 lista as fórmulas para calcular os diferentes recursos de textura do GLCM e as descrições correspondentes. Neste trabalho, os recursos estatísticos de textura, como contraste, correlação, dissimilaridade, energia, entropia, homogeneidade, média e variância, são calculados como os recursos de textura da imagem de couro fornecida.

Usando os recursos extraídos usando histograma de cores, autocorrelação e GLCM, um conjunto de dados rotulados é desenvolvido para treinar o estado dos classificadores de aprendizado de máquina, como rede neural de avanço raso (SFFNN), máquina de vetor de suporte (SVM), K- vizinho mais próximo (KNN) para classificação multiclasse de defeitos de couro.

4.2 Máquina baseada em rede neural de alimentação direta rasa

classificador de aprendizagem

Neste trabalho, uma rede neural feed-forward rasa (SFFNN) é treinada para classificar defeitos de couro, como marcas de dobras, grão de, pinhole, marcas de crescimento e grão solto. A Figura 7 mostra a arquitetura típica do SFFNN proposto com duas camadas ocultas. Aqui, os recursos de cor e o recurso de textura são usados como o vetor de entrada (x_i). Como a magnitude dos recursos de cor e textura extraídos é diferente, a normalização baseada na unidade é seguida para garantir a fusão adequada dos recursos extraídos para reduzir o viés e as influências grosseiras.

Também garante os valores do vetor de entrada no intervalo $[0, 1]$.

Uma rede neural feedforward rasa pode ser considerada como um modelo não linear com função de base não linear.

ções $\tilde{y}_j(x)$ dadas pela Eq. (12).

$$y(x,w) = f \left(\sum_{j=1}^H \tilde{y}_j^H w_{j\tilde{y}}(x) \right) \quad (12)$$

Aqui os pesos W_j podem ser ajustados durante o treinamento e $\tilde{y}_j(x)$ é uma função não linear de uma combinação linear de entradas. x refere-se ao vetor de recurso de cor e textura extraído da imagem de couro fornecida. A saída da rede neural feed-forward (y) pode ser expressa como uma série de transformações funcionais conforme dado pela Eq. (13).

$$z_k = g \left(\sum_{j=0}^H \tilde{y}_j^H w_{kj}^{(2)} h \left(\sum_{i=0}^D \tilde{y}_i^D w_{ji}^{(1)} x_i \right) \right) \quad (13)$$

Aqui, os sobrescritos (1) e (2) indicam os parâmetros das respectivas camadas ocultas, x_i indica o vetor de recursos de entrada. j refere-se ao número 'H' de nós ocultos. K refere-se ao número de neurônios de saída. g, h é a função de ativação não linear da camada oculta e utilizou a função de ativação sigmoide.

Para obter a classificação multiclasse do tipo de couro, como marcas de dobra, marcas de crescimento, grão, grão solto, orifício e couro não defeituoso, a função softmax é aplicada conforme indicado abaixo:

$$P(y = j|z(i)) = \frac{\tilde{y}(z(i))}{\sum_{j=0}^k \tilde{y}(z(i))} = \frac{e^{z(i)}}{\sum_{j=0}^k e^{z(j)}} \quad (14)$$

Essa função softmax calcula a probabilidade da amostra de treinamento dada $x(i)$ pertencer à classe j dado o peso e a entrada líquida $z(i)$. Portanto, calculamos a probabilidade $p(y=j|x(i); w_j)$ para cada rótulo de classe em $j=1, \dots, k$. Aqui, o termo de normalização no denominador faz com que essas probabilidades de classe sejam somadas a um.

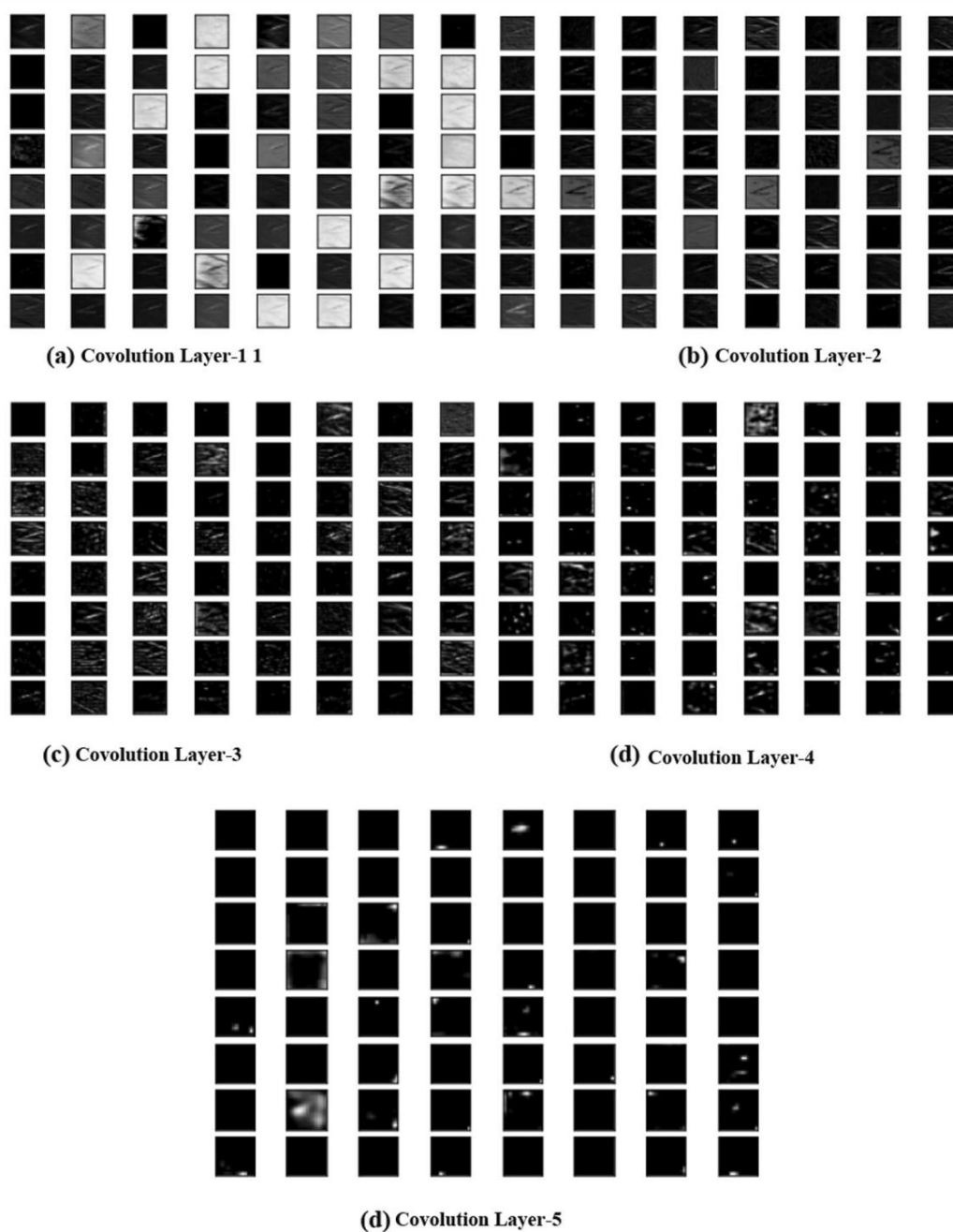


Fig. 8 Mapas de recursos típicos da Rede Neural Convolutacional Alexnet Deep Learning

Além disso, a rede neural rasa é treinada ajustando os pesos definindo e minimizando uma função de custo J , que é a média de todas as entropias cruzadas sobre o conjunto de dados de treinamento conforme abaixo:

$$J(W) = \frac{1}{n} \sum_{i=0}^n H(T_i, O_i) \quad (15)$$

Aqui, a função (H) refere-se à função de entropia cruzada conforme definido abaixo

$$H(T_i, O_i) = - \sum_m T_i \cdot \log(O_i) \quad (16)$$

Aqui, o T corresponde aos rótulos de “alvo” e o O representa a probabilidade calculada da função SoftMax. A função de custo baseada em entropia cruzada é

minimizado para o dado conjunto de dados de treinamento usando o método de descida de gradiente estocástico atualizando iterativamente a matriz de peso até que o número especificado de épocas ou limite de custo desejado seja atingido.

5 Métricas de desempenho de classificadores de aprendizado profundo e aprendizado de máquina

Neste trabalho, a matriz de confusão é construída para avaliar o desempenho de classificação das redes neurais de aprendizado profundo e classificadores de aprendizado de máquina. Uma matriz de confusão é uma matriz $N \times N$ onde N é o número de classes alvo e compara o número de classes preditas pelo classificador. Aqui, os elementos diagonais representam as classificações corretas, enquanto todas as outras entradas mostram classificações incorretas.

Com base na matriz de confusão, métricas de desempenho como precisão, sensibilidade, pontuação F1 e precisão são calculadas usando a fórmula fornecida abaixo.

$$\text{Precisão} = \frac{PT}{TP + FP} \quad (17)$$

$$\text{Sensibilidade} = \frac{PT}{PT + FN} \quad (18)$$

$$\text{Pontuação F1} = \frac{2 \text{ Sensibilidade} \times \text{Precisão}}{\text{Sensibilidade} + \text{Precisão}} \quad (19)$$

$$\text{Precisão geral} = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

onde TP (verdadeiro positivo) são as amostras de couro positivas classificadas corretamente, TN (verdadeiro negativo) são as amostras de couro negativas corretamente classificadas FP (falsos positivos) são as amostras negativas classificadas incorretamente e FN (falso negativo) são as amostras de couro positivas classificadas incorretamente amostras.

6 Resultados e discussão

Neste trabalho, um sistema de computação de aprendizado profundo envolvendo o sistema Windows 10 de 64 bits com CPU i5, 16 GB de memória e frequência básica de 2,30 GHz é usado para desenvolver redes neurais de aprendizado profundo e modelos de aprendizado de máquina para classificação multiclasse de texturas de couro.

Modelos propostos de redes neurais de aprendizado profundo e modelos de aprendizado de máquina usando as ferramentas de programação com o software MATLAB (Versão.2021b). O desempenho de treinamento e teste de redes neurais de aprendizado profundo pré-treinadas populares, como Alexnet, VGG-16, Google net, Squeeze Net, ResNet-50, são avaliados para a classificação de marcas de dobra, granulação, pinhole, marcas de crescimento, granulação solta e defeitos de couro não defeituosos. Mais longe,

a precisão da classificação de redes neurais convolucionais de aprendizado profundo é comparada com as abordagens de aprendizado de máquina existentes, como rede neural de avanço raso, máquina de vetor de suporte e K vizinho mais próximo. As métricas de desempenho de classificação, como precisão, sensibilidade, pontuação f1 e exatidão, são calculadas usando a matriz de confusão e os resultados são apresentados nesta seção.

6.1 Mapas de recursos de redes neurais de convolução

A fim de entender e interpretar os mapas de recursos nas camadas de convolução do Alexnet, VGG-16, Google net, Squeeze Net, ResNet-5, ele é extraído para algumas camadas de convolução. A Figura 8 mostra mapas de recursos de uma imagem de couro no Alexnet para as 5 camadas de convolução. Pode ser visto na Fig. 8a que características simples como arestas são filtradas por kernels na primeira camada de convolução e características de alta ordem são extraídas nas camadas subsequentes usando os pesos aprendidos dos kernels como mostrado na Fig. 8b-e.

À medida que a profundidade da camada aumenta, os mapas de recursos não mostram muitos detalhes devido ao tamanho mais fino do filtro no mesmo campo receptivo. Embora seja difícil interpretar os mapas de recursos extraídos, ele fornece recursos de baixo e alto nível de variações de textura de couro na imagem de couro no mesmo campo receptivo que é usado para classificação de imagens de couro.

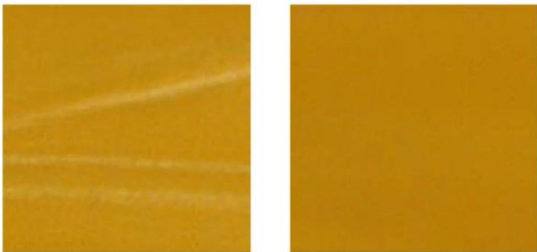
6.2 Extração de características usando GLCM, autocorrelação

Histograma de cores, GLCM, funções de autocorrelação são aplicadas para extrair as cores, recursos de textura das imagens de couro defeituosas e não defeituosas para avaliar o desempenho das abordagens de aprendizado de máquina, como para classificação de defeitos de couro multiclasse. A Figura 9 mostra os resultados da extração de características de couro defeituoso típico com uma marca de dobra e uma imagem de couro não defeituoso de cor amarela. Como a presença de defeitos no couro resulta em variações de intensidade, há uma variação no número de contagens do canal vermelho, verde, azul conforme mostra a Fig. 9b.

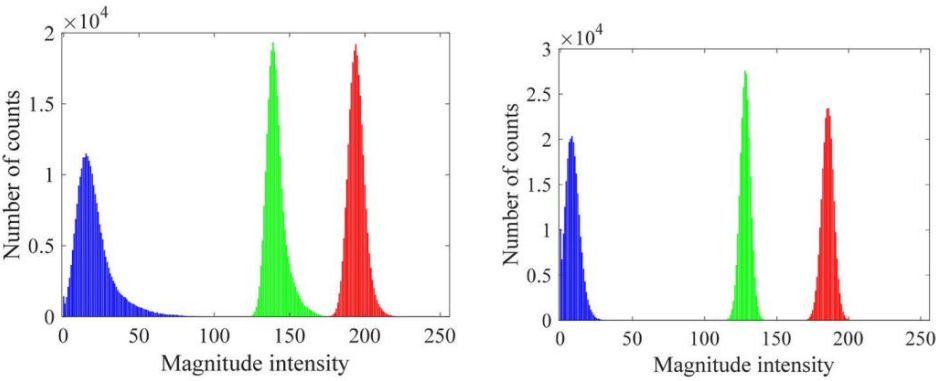
A partir da Fig. 9c, pode-se notar que, portanto, a função de autocorrelação do couro defeituoso com marcas de dobramento decai lentamente devido à textura mais grossa em comparação com o couro não defeituoso com textura mais fina, conforme mostrado na

Figura 9c.

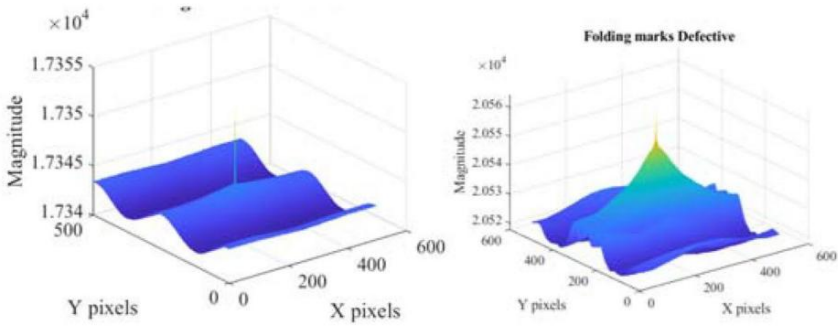
Como existe um padrão de textura diferente em couro defeituoso e não defeituoso, isso resulta em variações nas variações de intensidade e nos níveis de cinza dos pixels vizinhos, o que leva a uma mudança na magnitude de GLCM, conforme mostrado na Fig. 9d. A partir desses resultados, nota-se que as características de cor e textura extraídas variam devido ao padrão de textura diferente dos defeitos do couro e às variações de intensidade correspondentes.



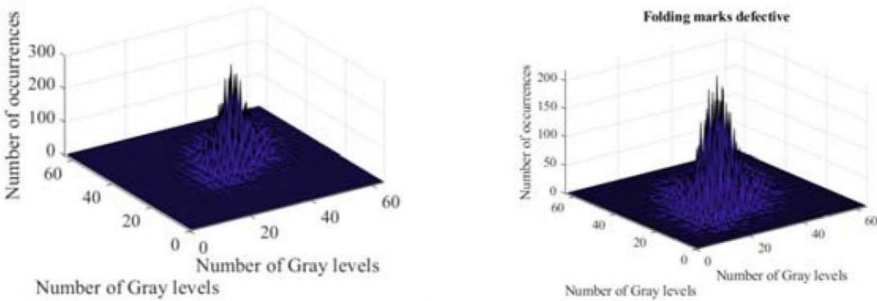
(a) Defective leather with folding mark and non-defective leather image



(b) Color histogram



(c) Autocorrelation



(d) GLCM

Fig. 9 Extração de características de cor e textura para couro defeituoso e não defeituoso

Tabela 3 Parâmetros da descida do gradiente estocástico com momento (SGDM)

Hiperparâmetros	Valor
Algoritmo de otimização	SGDM
Taxa de aprendizado inicial	0,0001
épocas	220
taxa de lote	30

6.3 Treinando e testando o desempenho do aprendizado profundo redes neurais

Neste trabalho, a Descida de Gradiente Estocástico com momento (SGD) é usada para otimizar os hiperparâmetros do modelo, particularmente a taxa inicial, passo, tamanho do filtro de profundidade

aprendendo redes neurais. A Tabela 3 mostra os detalhes do treinamento da rede neural.

A Figura 10 mostra as curvas de desempenho de treinamento e teste para as diferentes épocas e iterações do AlexNet. O gráfico inferior mostra a função de perda de entropia cruzada para diferentes épocas do conjunto de dados de treinamento (azul) e teste (preto). O gráfico superior mostra a tendência e variação da precisão de classificação do AlexNet ao longo das épocas.

A partir das parcelas, nota-se que o processo de treinamento convergiu bem para atingir a precisão de classificação de 99,4%. Pode-se ver que a magnitude da perda diminui para cada época com o aumento da precisão da classificação. A função de precisão e perda mostra solavancos à medida que os pesos das redes neurais são aprendidos a partir dos exemplos dados de treinamento e teste de imagens de couro para classificação multiclasse. o decorrido

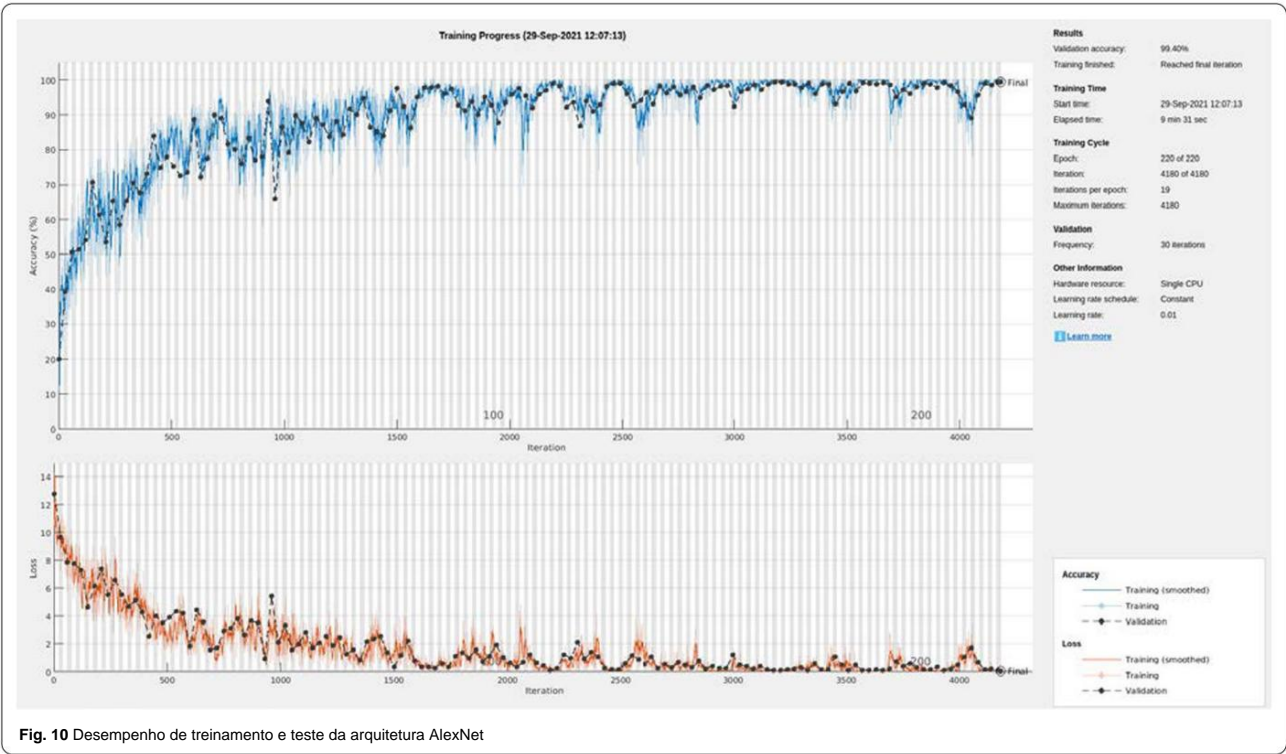
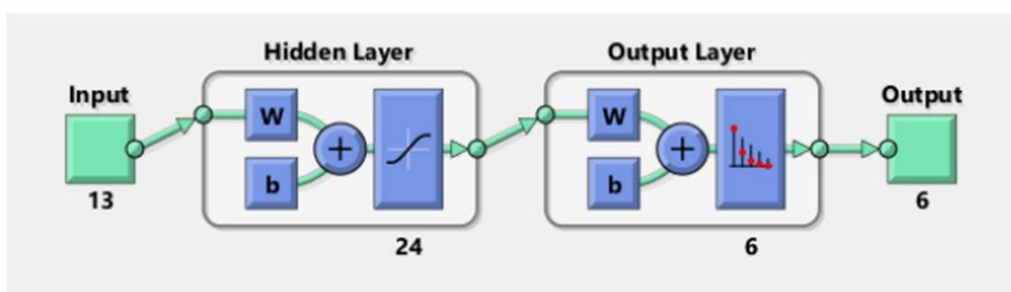


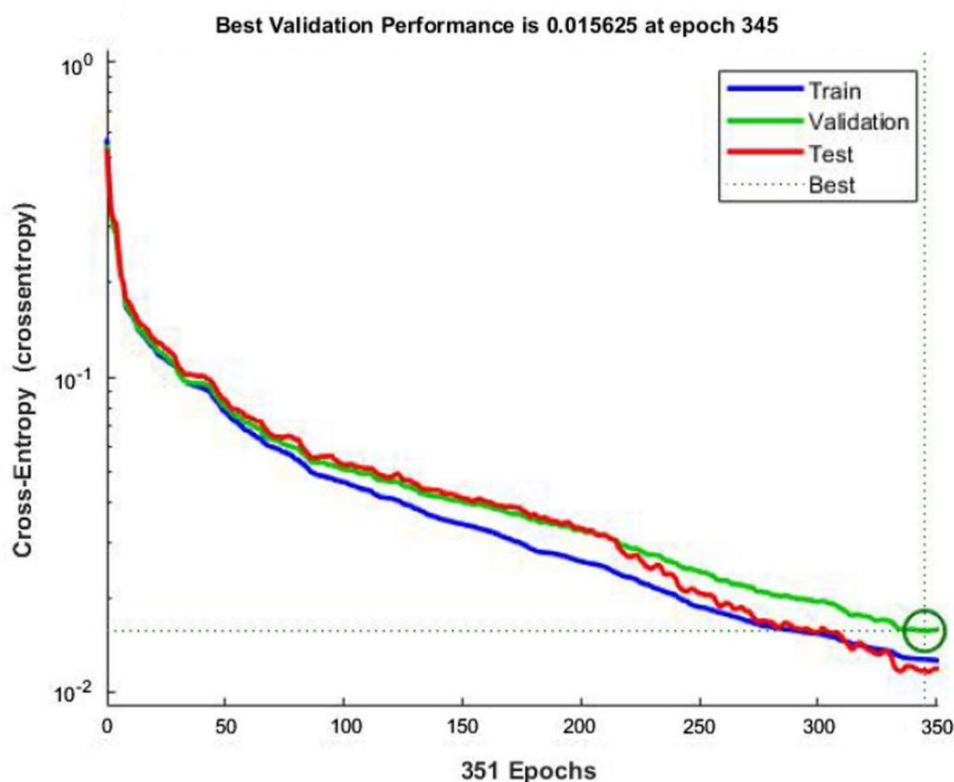
Fig. 10 Desempenho de treinamento e teste da arquitetura AlexNet

Tabela 4 Tempo de treinamento computacional do modelo de aprendizagem profunda Alexnet

Época	Iteração	Tempo decorrido (hh:mm:ss)	Precisão de mini-lote (%)	Precisão de validação (%)	Perda de mini-lote	perda de validação	Taxa de aprendizagem básica
1	1	00:00:03	25.78	21.15	10.2304	12.5705	1.000eý04
12	210	00:00:31	60.16	72.98	6.2358	4.3037	1.000eý04
21	390	00:00:55	82.03	73.17	2.7172	4.2595	1.000eý04
124	2340	00:05:18	96.88	94.60	0.4982	0.8561	1.000eý04
220	4180	00:09:31	99.40	97.49	0.336	0.2404	1.000eý04



(a) Neural network architecture of leather defect classification



(b) Training and validation performance

Fig. 11 Configuração e treinamento do SFFNN

tempo computacional do processo de treinamento e a melhoria da precisão, a redução da perda é observada para diferentes números de épocas e é mostrada na Tabela 4.

Pode-se notar que a perda de minilotes diminui à medida que os valores de peso são aprendidos para a classificação correta com o aumento do número de iterações e a precisão da classificação chega a 99,40% em 220 Epoch.

6.4 Desempenho de treinamento de feed forward neural raso classificador de rede

Neste trabalho, uma arquitetura de rede neural rasa com 13 neurônios na camada de entrada, 24 neurônios na camada oculta

são desenvolvidos e treinados em ambiente MATLAB A Fig. 11 mostra a arquitetura da rede neural proposta e a camada de saída é configurada com 6 neurônios para classificação do tipo de imagem de couro com marcas de dobra, grão de, pinhole, marcas de crescimento, grão solto e superfície de couro não defeituosa. O treinamento e o desenvolvimento da rede neural feed-forward rasa proposta são realizados, uma função de entropia cruzada e um método de gradiente descendente são usados para ajustar os valores de peso da rede neural.

O gráfico de desempenho de treinamento, teste e validação do SFFNN proposto e o número de épocas é

Training accuracy

Output Class	Folding Marks	420 16.7%	0 0.0%	0 0.0%	5 0.2%	0 0.0%	2 0.1%	98.4% 1.6%
	Grain off	0 0.0%	420 16.7%	13 0.5%	5 0.2%	12 0.5%	0 0.0%	93.3% 6.7%
	Growth mark	0 0.0%	0 0.0%	394 15.6%	3 0.1%	2 0.1%	0 0.0%	98.7% 1.3%
	Loose grain	0 0.0%	0 0.0%	2 0.1%	407 16.2%	0 0.0%	0 0.0%	99.5% 0.5%
	Non defective	0 0.0%	0 0.0%	10 0.4%	0 0.0%	405 16.1%	3 0.1%	96.9% 3.1%
	pinhole	0 0.0%	0 0.0%	1 0.0%	0 0.0%	1 0.0%	415 16.5%	99.5% 0.5%
		100% 0.0%	100% 0.0%	93.8% 6.2%	96.9% 3.1%	96.4% 3.6%	98.8% 1.2%	97.7% 2.3%
		Target Class						
		Folding Marks	Grain off	Growth mark	Loose grain	Non defective	pinhole	

(a) Alexnet

Training accuracy

Output Class	Folding Marks	420 16.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	Grain off	0 0.0%	418 16.6%	0 0.0%	1 0.0%	4 0.2%	1 0.0%	98.6% 1.4%
	Growth mark	0 0.0%	0 0.0%	418 16.6%	0 0.0%	1 0.0%	0 0.0%	99.8% 0.2%
	Loose grain	0 0.0%	0 0.0%	0 0.0%	418 16.6%	0 0.0%	0 0.0%	100% 0.0%
	Non defective	0 0.0%	1 0.0%	0 0.0%	0 0.0%	412 16.3%	0 0.0%	99.8% 0.2%
	pinhole	0 0.0%	1 0.0%	2 0.1%	1 0.0%	3 0.1%	419 16.6%	98.4% 1.6%
		100% 0.0%	99.5% 0.5%	99.5% 0.5%	99.5% 0.5%	98.1% 1.9%	99.8% 0.2%	99.4% 0.6%
		Target Class						
		Folding Marks	Grain off	Growth mark	Loose grain	Non defective	pinhole	

(b) VGG 16

Training accuracy

Output Class	Folding Marks	390 15.5%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	8 0.3%	97.7% 2.3%
	Grain off	0 0.0%	420 16.7%	0 0.0%	2 0.1%	3 0.1%	4 0.2%	97.9% 2.1%
	Growth mark	0 0.0%	0 0.0%	419 16.6%	13 0.5%	55 2.2%	0 0.0%	86.0% 14.0%
	Loose grain	30 1.2%	0 0.0%	0 0.0%	396 15.7%	32 1.3%	3 0.1%	85.9% 14.1%
	Non defective	0 0.0%	0 0.0%	0 0.0%	8 0.3%	329 13.1%	13 0.5%	94.0% 6.0%
	pinhole	0 0.0%	0 0.0%	1 0.0%	1 0.0%	0 0.0%	392 15.6%	99.5% 0.5%
		92.9% 7.1%	100% 0.0%	99.8% 0.2%	94.3% 5.7%	78.3% 21.7%	93.3% 6.7%	93.1% 6.9%
		Target Class						
		Folding Marks	Grain off	Growth mark	Loose grain	Non defective	pinhole	

(c) ResNet 50

Training accuracy

Output Class	Folding Marks	382 15.2%	0 0.0%	0 0.0%	5 0.2%	0 0.0%	0 0.0%	98.7% 1.3%
	Grain off	0 0.0%	379 15.0%	0 0.0%	0 0.0%	24 1.0%	0 0.0%	94.0% 6.0%
	Growth mark	0 0.0%	0 0.0%	401 15.9%	0 0.0%	2 0.1%	1 0.0%	99.3% 0.7%
	Loose grain	5 0.2%	0 0.0%	14 0.6%	404 16.0%	114 4.5%	7 0.3%	74.3% 25.7%
	Non defective	33 1.3%	41 1.6%	0 0.0%	11 0.4%	272 10.8%	30 1.2%	70.3% 29.7%
	pinhole	0 0.0%	0 0.0%	5 0.2%	0 0.0%	8 0.3%	382 15.2%	96.7% 3.3%
		91.0% 9.0%	90.2% 9.8%	95.5% 4.5%	96.2% 3.8%	64.8% 35.2%	91.0% 9.0%	88.1% 11.9%
		Target Class						
		Folding Marks	Grain off	Growth mark	Loose grain	Non defective	pinhole	

(d) Inception v3

Fig. 12 Comparação da precisão de classificação dos classificadores de rede neural de convolução de aprendizagem profunda

mostrado na Fig. 11b. Verificou-se que a precisão de classificação de 97,6% para o valor mínimo de entropia cruzada de 0,015625 e melhor desempenho de validação é alcançado em 345 épocas.

6.5 Desempenho de classificação de redes neurais de aprendizado profundo

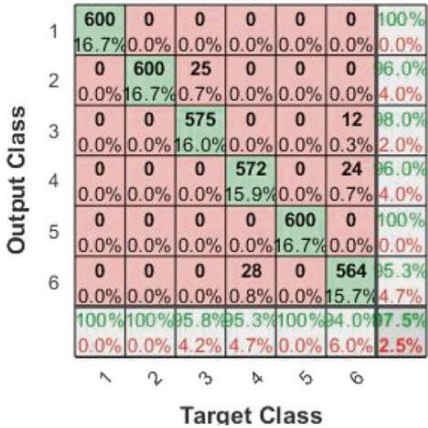
A fim de quantificar o desempenho de classificação das redes neurais de convolução de aprendizado profundo para o multi

classificação de classe de textura de couro, a matriz de confusão é calculada com base no número de classes de saída fornecidas pelo classificador e nas classes de destino fornecidas de imagens de couro. A Figura 12 mostra a matriz de confusão para redes neurais de aprendizado profundo e classificadores de aprendizado de máquina que forneceram as três maiores precisões de classificação durante o treinamento e o teste.

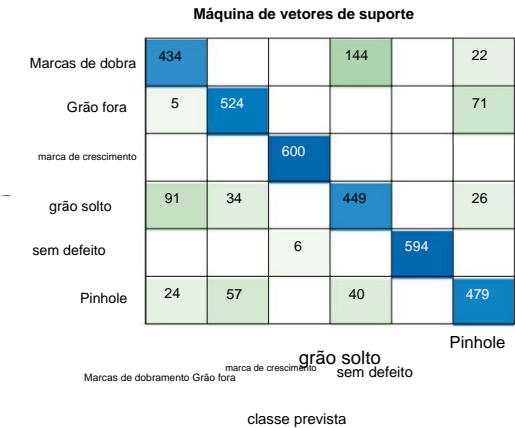
Na Fig. 12, os números altos nas células verdes representam respostas corretas e os números baixos nas células vermelhas correspondem

Tabela 5 Comparação da precisão da classificação de redes neurais de aprendizado profundo

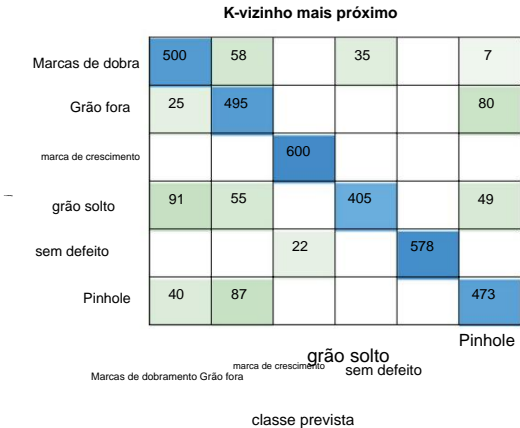
aulas de couro e precisão	AlexNet		VGG16		ResNet 50		A origem v3	
	Lembrar	Precisão	Lembrar	Precisão	Lembrar	Precisão	Lembrar	Precisão
Marcas de dobra	1	1	1	0,98	0,92	0,97	0,90	0,98
Grão de	0,99	0,97	1	0,93	0,99	0,97	0,90	0,94
marca de crescimento	0,99	1	0,93	0,98	0,94	0,86	0,95	0,99
grão solto	0,91	1	0,96	0,99	0,78	0,85	0,96	0,74
sem defeito	0,99	0,94	0,96	0,96	0,93	0,94	0,64	0,70
Pinhole	0,99	0,97	0,98	0,99	0,91	0,99	0,90	0,96
Precisão do treinamento	99,40%		97,7%		93,10%		88,1%	
Precisão de teste	87,60%		85,80%		83,60%		80,0%	



(a) Rede neural de alimentação rasa proposta



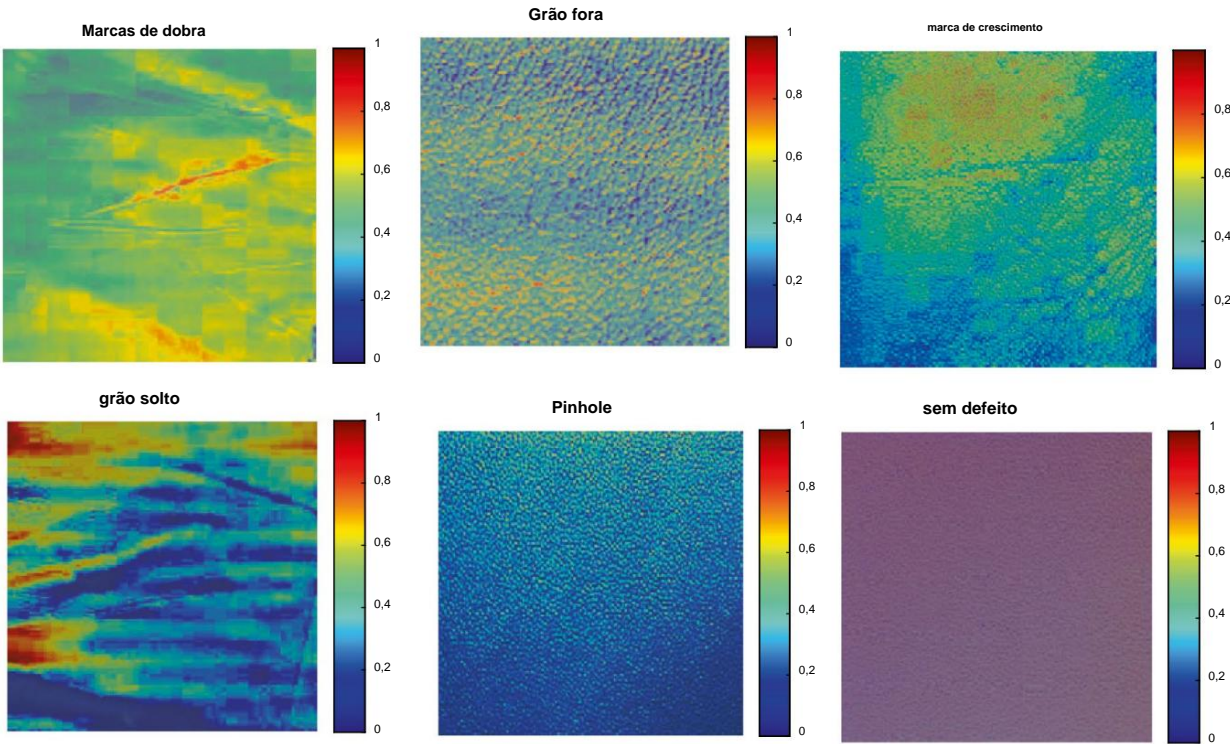
(b) Máquina de vetores de suporte



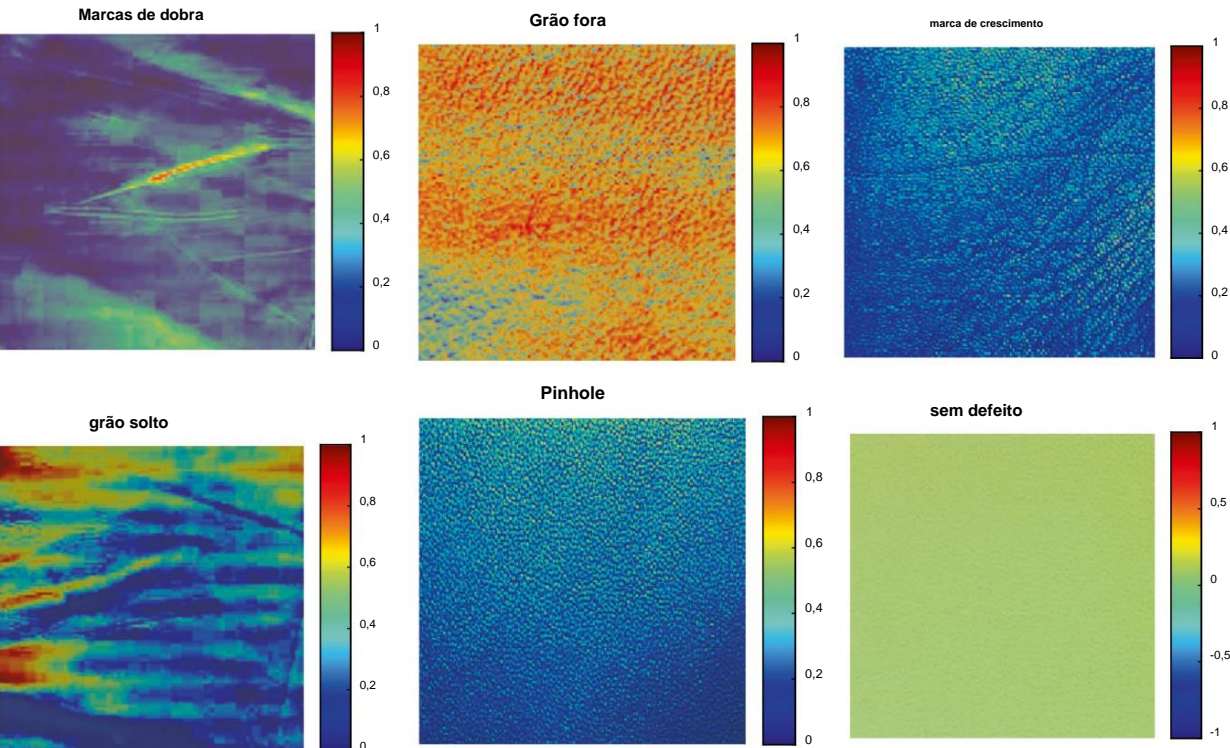
(c) K-Vizinho mais próximo

Fig. 13 Comparação da precisão de classificação dos classificadores de aprendizado de máquina

a respostas incorretas. Os valores percentuais na coluna da extrema direita da matriz de confusão mostram a precisão (ou valor preditivo positivo) e a taxa de descoberta falsa da rede neural proposta para a classificação de cada classe de couro. Além disso, a linha inferior da matriz de confusão mostra a sensibilidade e a taxa de falsos positivos. Devido à capacidade superior de extração de recursos com a estrutura do kernel, o Alexnet teve um desempenho melhor com

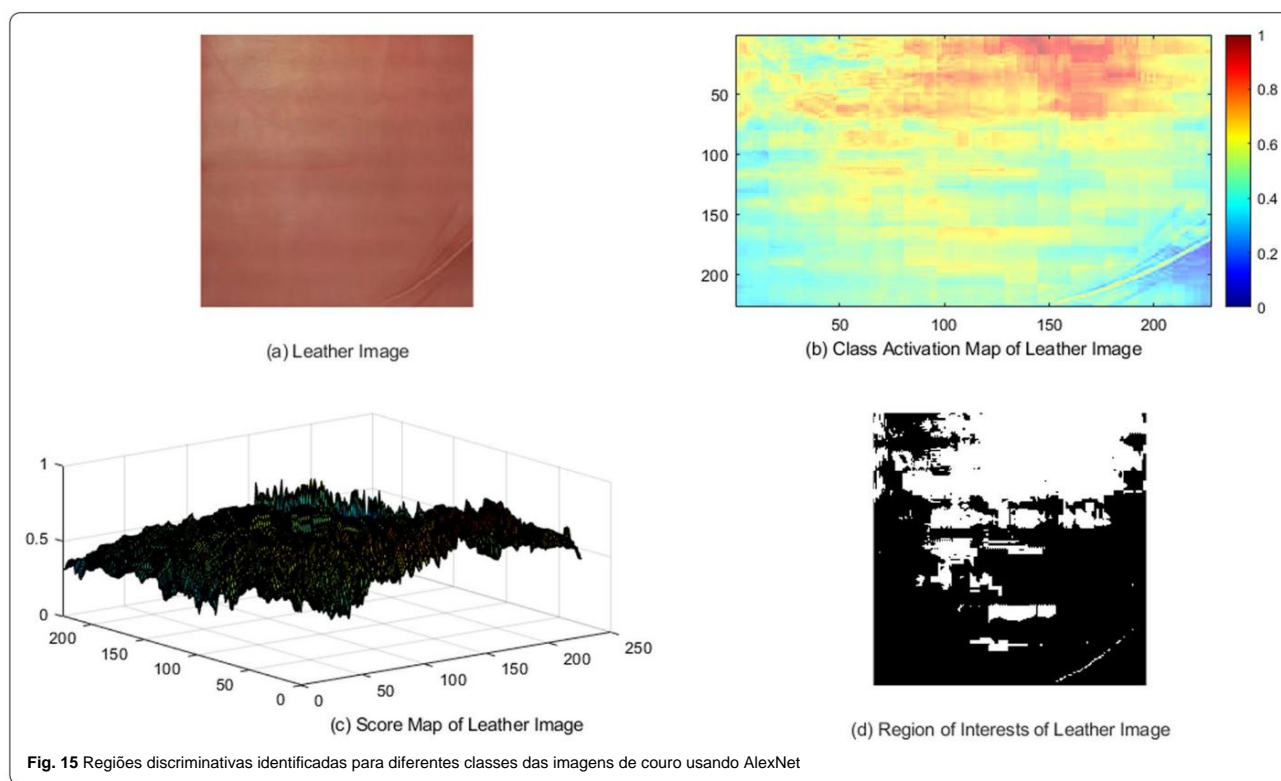


(a) AlexNet



(b) VGG16

Fig. 14 Regiões discriminativas identificadas para diferentes classes das imagens de couro usando redes neurais de aprendizado profundo



a precisão de classificação de 99,40% e 87,60% para os conjuntos de dados de treinamento e teste do que os outros classificadores de rede neural de convolução de aprendizagem profunda.

Usando o número de classificações corretas e incorretas da classe alvo das imagens de couro, a precisão, valores de recall são calculados e listados na Tabela 5 para diferentes classificadores.

Com a melhor precisão de teste de 87,60%, o Alexnet teve melhor desempenho para classificação multiclasse de texturas de couro de imagens de couro não vistas. Esses resultados provaram a capacidade de classificadores de rede neural de aprendizado profundo para a aplicação de classificação de textura de couro multiclasse.

6.6 Desempenho de classificação de abordagens de aprendizado de máquina

Algoritmos de aprendizado de máquina de última geração, como redes neurais de alimentação rasa, máquina de vetor de suporte, K vizinho mais próximo, árvore de decisão, Naïve Bayes são aplicados usando recursos de cor e textura artesanais para classificação de textura de couro multiclasse. A matriz de confusão é construída para resumir o desempenho dos diferentes classificadores de aprendizado de máquina e é mostrada na Fig. 13. Conforme indicado pela maior magnitude dos elementos diagonais na matriz de confusão para a classificação correta de cada classe, rede neural de avanço raso

apresentou melhor desempenho na classificação de diferentes imagens de textura de couro. A precisão geral da classificação da rede neural de avanço raso para a classificação de imagens de couro multiclasse é de 97,5%, o que é menor do que os modelos de rede neural de convolução de aprendizado profundo, pois a extração de recursos artesanais limita os recursos discriminativos importantes nas imagens de couro.

6.7 Mapas de ativação de classe para seleção de região de interesse em imagens de couro

Como os defeitos de couro estão localizados nas regiões específicas da imagem de couro, os mapas de ativação de classe são gerados usando os modelos de rede neural de aprendizado profundo treinados para as imagens de couro dadas e os resultados da amostra para diferentes classes de textura de couro são mostrados na Fig. 14a, b para Alexnet e VGG-16, respectivamente.

Pode-se ver que as regiões com a cor vermelha são identificadas como as regiões discriminativas e podem ser identificadas a partir dos pixels com maior magnitude usando o mapa de pontuação. Neste trabalho, o valor máximo do mapa de ativação da classe é escolhido e aplicado como limiar para segmentação da região de interesse nas imagens de couro. Figura 15a mostra de imagem de couro com caixa delimitadora de verdade nas áreas defeituosas e o mapa de ativação de classe correspondente na Figura 15b. Pode-se ver que o mapa de pontuação mostra picos e valores mais altos no

respectivas regiões da imagem identificadas em vermelho na Fig. 15b.

Um valor limite de 0,5983 é selecionado para segmentar a região de interesse e os resultados são mostrados na Fig. 15c. Verificou-se que a área de regiões discriminativas no couro identificadas pelo mapa de ativação de classe é maior do que a caixa delimitadora de verdade de campo. Portanto, requer algoritmos adequados para a detecção precisa e localização de defeitos de couro nas imagens de couro.

7 Conclusões

A textura do couro desempenha um papel importante na decisão da qualidade dos produtos de couro. Este trabalho apresentou redes neurais convolucionais de aprendizado profundo e classificadores de aprendizado de máquina para as imagens de couro de classificação multiclasse. Um conjunto de dados abrangente de 3.600 imagens de couro com diferentes defeitos, como marcas de dobra, grão, orifício, marcas de crescimento, grão solto e superfícies de couro não defeituosas são classificadas usando redes neurais de aprendizado profundo pré-treinadas, como Alexnet, VGG-16, Google net, Squeeze Net, ResNet-50. O desempenho da classificação de redes neurais convolucionais de aprendizado profundo é comparado com as abordagens de aprendizado de máquina existentes, como rede neural de avanço raso, máquina de vetor de suporte e K-vizinho mais próximo. A partir dos resultados obtidos a partir da matriz de confusão, descobriu-se que a rede neural de convolução de aprendizado profundo como Alexnet teve um desempenho melhor com a precisão de classificação de 99,4% do que a técnica de aprendizado de máquina de rede neural de avanço raso devido à capacidade superior de extração de recursos. Além disso, o uso de mapas de ativação de classe da rede neural de aprendizagem profunda treinada para segmentar as regiões de interesse nas imagens de couro é demonstrado para a localização das regiões defeituosas. O método proposto pode ser adequadamente implementado para classificação multiclasse automatizada de amostras de couro em um ambiente industrial.

Reconhecimentos

Os autores agradecem à administração do VIT por fornecer as instalações experimentais necessárias para realizar pesquisas em inspeção de couro. Os autores também agradecem à KH Leather Company, Vellore, por fornecer amostras de couro para inspeções.

Contribuições dos autores

A PKM realizou o processamento de imagem e desenvolvimento de algoritmo para classificação de defeitos de couro. O DASS forneceu a estrutura para o modelo de rede neural para treinar e testar a aplicação da classificação de defeitos de couro. Ambos os autores leram e aprovaram o manuscrito final.

Financiamento

Não aplicável.

Disponibilidade de dados e materiais

Um conjunto de dados de imagem de couro contendo diferentes defeitos, como marcas de dobra, granulação, pinhole, marcas de crescimento, granulação solta e couro não defeituoso

superfícies está disponível em Kaggle: <https://www.kaggle.com/praveen2084/leather-defect-classification>.

Declarações

Aprovação ética e consentimento para participar

Não aplicável.

Consentimento para publicação

Não aplicável.

Interesses competitivos

Os autores declaram que não têm interesses concorrentes.

Recebido: 15 de março de 2021 Aceito: 21 de janeiro de 2022

Published online: 16 March 2022

Referências

1. Lovergine FP, Branca A, Attolico G, Distância A. Inspeção de couro por análise de textura orientada com uma abordagem morfológica. In: Proceedings of International Conference Image Processing, vol 2. 1997. p. 669–671.
2. Yeh C, Perng DB. Um padrão de referência de compensação de defeitos para transações de couro. Int J Adv Manuf Technol. 2005;25(11–12):1197–204.
3. Newman TS, Jain AK. Levantamento de inspeção visual automatizada. Computação Vis Image Underst. 1995;61:231–62.
4. Georgiou L, Krastev K, Angelov N. Identificação de defeitos de superfície de couro. In: Proceedings of the 4th International Conference on Computer Systems and Technologies e-learning. 2003. pág. 303–307.
5. Kavitha C, Karthika K, Hamsaveni VG, Supriya B, Manasa BMR. Identificação e classificação de defeitos em diferentes etapas do processamento do couro usando sistema baseado em visão. Int J Adv Sci Technol. 2019;28(19):123–8.
6. Mihan T, Anil J. Análise de textura. In: Chen CH, Pau LF, Wang PSP, editores. O manual de reconhecimento de padrões e visão computacional. Cingapura: World Scientific Publishing Co; 1993. pág. 235–276.
7. Jawahar M, Chandra Babu NK, Vani K. Sistema de inspeção de visão de máquina para detecção de defeitos na superfície do couro. J Am Leather Chem Assoc. 2019;114(1):10–19.
8. Sobral JL. Inspeção de couro baseada em wavelets. In: Marques JS, de la Blanca NP, Pina P, editores. Reconhecimento de padrões e análise de imagens. Berlim: Springer; 2005. pág. 682–8.
9. Branca A, Lovergine FP, Attolico G, Distância A. Detecção de defeitos em couro por singularidades orientadas. In: Conferência Internacional sobre Análise Computacional de Imagens e Padrões. Berlim: Springer; 1997. pág. 223–230.
10. Aslam M, Khan TM, Naqvi SS, Holmes G, Nafa R. Sobre a aplicação de visão de máquina automatizada para inspeção e classificação de defeitos de couro: uma pesquisa. Acesso IEEE. 2019;7:176065–86.
11. Haralick RM, Shanmugam K, Dinstein I. Recursos de textura para classificação de imagens. IEEE Trans Syst Man Cybern. 1973;6:610–21.
12. Jawahar M, Babu NKC, Vani K. Classificação da textura do couro usando a técnica de extração de características wavelet. In: Anais da conferência internacional IEEE sobre inteligência computacional e pesquisa em computação. 2014. pág. 1–4.
13. Hemerson P, Willian PA, Priscila SM, Mauro CP, Pereira MA, Jacinto MAC. Detecção de defeitos em couro cru e couro wet blue. In: ComplIMAGE, modelação computacional de objectos representados em imagens: fundamentos, métodos e aplicações, Coimbra. 2007. pág. 355–360.
14. Viana R, Rodrigues RB, Alvarez MA, Pistori H. SVM com seleção estocástica de parâmetros para classificação de defeitos em couro bovino. In: Mery D, Rueda L, editores. Avanços na tecnologia de imagem e vídeo. PSIVT 2007. Notas de aula em ciência da computação, vol. 4872. Berlim: Springer; 2007.
15. Kwak C, Ventura JA, Tofang-Sazi K. Inspeção automatizada de defeitos e classificação de tecido de couro. Análise de dados Intell. 2001;5(4):355–70.
16. Jawahar M, Babu NKC, Vani K, et al. Sistema de inspeção baseado em visão para detecção de defeitos de superfície de couro usando abordagem de classificação de conjunto de otimização de enxame de partículas de convergência rápida. Multimed Tools Appl. 2021;80:4203–35.

17. Villar P, Mora M, Gonzalez P. Uma nova abordagem para a segmentação de defeitos de couro wet blue. In: San Martin C, Kim SW, editores. CIARP 2011, LNCS 7042. Berlin: Springer; 2011. pág. 591–8.
18. Amorim WP, Pistori H, Pereira MC, Jacinto MAC. Redução de atributos aplicada à classificação de defeitos de couro. In: Anais do 23º Congresso SIBGRAP sobre gráficos, padrões e imagens. 2010. pp. 353–359.
19. He K, Zhang X, Ren S, Sun J. Aprendizado residual profundo para reconhecimento de imagem. 2015. arXiv: 0706.1234.
20. Ren R, Hung T, Tan KC. Uma abordagem genérica baseada em aprendizado profundo para inspeção automatizada de superfícies. IEEE TransCybern. 2018;48(3):929–40.
21. Liong S, Gan YS, Huang Y, Liu K, Yau W. Rede neural integrada e abordagem de visão de máquina para classificação de defeitos de couro. 2019. arXiv: 1905.11731.
22. Liong ST, Zheng D, Huang YC, Gan Y. Classificação e segmentação de defeitos de couro usando arquitetura de aprendizado profundo. Int J Comput Integr Manuf. 2020;33(10):1–13.
23. Xu D, Wang CS, Wu CF. O reconhecimento de características de cor e textura baseado nos sentidos da visão e do tato. In: Anais da 13ª Conferência Internacional sobre Linguagem, Educação, Humanidades e Inovação e 2ª Conferência Internacional sobre Aprendizagem Aberta e Tecnologias Educacionais. 5 e 6 de abril de 2019. 2019. p. 22–29.
24. Alzubaidi L, Zhang J, Humaidi AJ, et al. Revisão do aprendizado profundo: com conceitos, arquiteturas CNN, desafios, aplicações, direções futuras. J Grandes Dados. 2021;8:53. <https://doi.org/10.1186/s40537-021-00444-8>.

Nota do editor

A Springer Nature permanece neutra em relação a reivindicações jurisdicionais em mapas publicados e afiliações institucionais.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
