

COMPARAÇÃO ENTRE MÉTODOS DE CLASSIFICAÇÃO SUPERVISIONADA

CLASSIFICAÇÃO DIGITAL DE IMAGENS

GONÇALVES 1, Rômulo P. 1; ASSIS 2, Leonardo C. 2; ANDRADE 3, Rafael J.O. 3; MAROTTA 4, Giuliano S. 4.

PALAVRAS CHAVE: Classificação Supervisionada; Sensoriamento Remoto; Algoritmos.

RESUMO

Para atender às necessidades do usuário final que, muitas vezes não tem o conhecimento necessário para interpretar uma imagem no seu estado bruto (numérico) faz-se necessário transformá-la numa imagem temática que mostre apenas as informações de seu interesse. O resultado final pode ser obtido, em geral, por dois tipos de classificação: a classificação supervisionada que exige um conhecimento *a priori* da área estudada e a classificação não supervisionada que dispensa essas informações iniciais.

Os materiais utilizados foram uma imagem do sensor TM da plataforma orbital LandSat7, composta pelas 7 bandas disponíveis do sensor e o software Idrisi 3.2 para efetuar a aplicação dos algoritmos de classificação.

A classificação de imagens digitais associa a cada pixel da imagem um rótulo ou um tema de interesse, dessa forma uma imagem numérica é transformada numa imagem temática onde o número de categorias é definido pelo analista. São abordados dois tipos de classificação supervisionada, sendo que a primeira consta da utilização do algoritmo do paralelepípedo, e a segunda do algoritmo da verossimilhança.

Ao submetermos as imagens brutas neste processo, o software possibilita como geração de resultados, uma matriz de contingência ou de erros onde se pode avaliar a qualidade da classificação nos dois algoritmos.

O algoritmo do paralelepípedo utiliza a média e o desvio padrão, ou os valores máximos e mínimos das amostras de treinamento. Um pixel é atribuído à uma dada classe se satisfizer a condição apresentada na Equação 01.

$$\mu_{ck} - S_{ck} \leq VN_{ijk} \leq \mu_{ck} + S_{ck} \quad (01)$$

onde,

$c = 1, 2, 3, \dots, m \rightarrow n^\circ$ de classes;

$k = 1, 2, 3, \dots, l \rightarrow$ n° de bandas;

μ_{ck} = média da classe c na banda k ;

S_{ck} = desvio-padrão da classe c na banda k ;

V_{nik} =Valor numérico do pixel desconhecido na linha i , coluna j , banda k .

No algoritmo da verossimilhança, utiliza-se a média e a covariância das amostras de treinamento, para um cálculo estatístico da probabilidade de um pixel pertencer a uma outra classe, utilizando funções de densidade de probabilidade. Depois de avaliar a probabilidade para cada categoria, o pixel é assinalado à classe cuja probabilidade tenha sido a mais alta, ou nenhuma delas no caso de os valores de probabilidade estarem abaixo de um determinado limite. Como critério de decisão, foi abordada a condição apresentada na Equação 02.

$$P_c \geq P_i \quad (02)$$

onde,

$i = 1, 2, 3 \dots m$ possíveis classes;

P_c = probabilidade do vetor X ser assinalado na classe c ;

P_i = probabilidade do vetor X ser assinalado a qualquer outra classe;

X = vetor das medidas dos pixels desconhecidos onde n é o número de bandas (Equação 03).

$$X^T = [VN_{ij1} \quad VN_{ij2} \quad VN_{ij3} \quad \dots \quad VN_{ijn}] \quad (03)$$

A função de densidade de probabilidade (P_c) é apresentada pela Equação 04:

$$P_c = \left[-0.5 \log_e (\det(V_c)) \right] - \left[0.5 (X - M_c)^T (V_c)^{-1} (X - M_c) \right] \quad (04)$$

onde,

$\det(V_c)$ = é o determinante da matriz de covariância V_c ;

M_c = valor das médias para cada classe c ;

V_c = matriz de covariância da classe c contemplando todas as bandas.

Foram determinadas e numeradas oito classes de usos do solo, conforme apresentado na Tabela 01.

Tabela 01: Classes de uso e cobertura vegetal

Classe	Classificação
1	Vegetação Madura
2	Vegetação Jovem
3	Solo Seco
4	Solo Úmido
5	Lago Limpo
6	Lago Turvo
7	Brejo
8	Rio

De acordo com as matrizes calculadas, os resultados dos algoritmos de máxima verossimilhança e distância mínima até a média, as classes 3, 4, 6 e 8 não apresentaram erro de omissão em ambos, apenas na classe 1, o de máx. verossimilhança teve um erro de omissão maior.

Analisando o erro de comissão percebe-se que as classes 5, 6 e 8, não apresentaram-no em ambos algoritmos, e já a classe 3, não apresenta este erro, apenas no de algoritmo de máxima verossimilhança. Somente na classe 2, o último algoritmo apresentou erro de comissão maior.

Os índices Kappa alcançados pelos algoritmos de máxima verossimilhança e distância mínima até a média foram 0,8506 e 0,8206 respectivamente.

Conclui-se que a classificação de imagens feitas pelo método de classificação supervisionada utilizando os algoritmos da distância mínima até a média e máxima verossimilhança, apresentaram resultados satisfatórios para a área em estudo, porém o que apresentou o melhor resultado foi o algoritmo de verossimilhança.

Em cada caso, é necessário que o analista estude os diferentes resultados apresentados pelos diversos algoritmos disponíveis no software, optando assim pelo de melhor resultado, pois a qualidade dos resultados pode variar de acordo com a região analisada.

*UFV – Universidade Federal de Viçosa
Rua dos Estudantes, 100/5B, Centro
CEP: 36570-000 - VIÇOSA – MG
BRASIL*

*Tel: (0 + 31) 3892 3002
Fax: (0 + 31) 8713 3675
romuloparma@yahoo.com.br
Engenheiro Agrimensor
Mestrado em Informações Espaciais*