

Classificação de Imagens Utilizando Fusão de Sensores Termal e Visível

Raoni Avilez Fiedler¹ e Francisco Bernardo Lovato Eick¹

¹Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP – Brasil

Resumo – Utilizando uma câmera com sensor duplo (visível e termal), este trabalho avalia a alteração na exatidão global de quatro classes de interesse utilizando-se diferentes composições de canais nas imagens analisadas. São testadas as composições RGB e RGBI (composição RGB mais canal infravermelho). Os resultados são comparados utilizando os algoritmos k vizinhos mais próximos (k-NN) e máquina de vetores de suporte (SVM). Os resultados experimentais indicam que o uso da composição RGBI aumenta a acurácia na classificação em 9,7%, no k-NN, e 1,9% no SVM.

Palavras-Chave – sensoriamento remoto, infravermelho, reconhecimento de padrões.

I. INTRODUÇÃO

No âmbito do Sensoriamento Remoto, a utilização de sensores ópticos com operação fora da faixa visível do espectro eletromagnético estende a capacidade humana de obter informações visuais sobre aquilo que muitas vezes não ficaria evidente em uma imagem convencional.

Sensores com faixa de operação no infravermelho permitem obter informações a partir do diferencial de temperatura de cada objeto imageado na cena. A sua aplicação encontra-se consolidada tanto nas áreas de defesa quanto civil. É crescente a aplicação desse tipo de sensor em drones em tarefas como acompanhamento de urbanização [1] e plantio [2] ou de painéis solares [3] assim como missões de busca e salvamento em terra [4] e em mar [5], [6], podem ser feitas de maneira flexível, sem limitações como a baixa resolução espacial ou temporal de uma imagem satelital.

Sensores com diversos canais podem extrair camadas adicionais de dados, permitindo uma visualização mais completa do ambiente de interesse como visto em [7], [8]. A extração de camadas adicionais de dados do cenário de operação representa uma vantagem para as aplicações em área de defesa [9].

Nesse sentido, torna-se interessante uma análise da viabilidade da fusão de dados dos sensores disponíveis de modo a correlacionar as informações obtidas pelos diversos canais. Essa técnica pode ser empregada por drones [10]-[12], aeronaves tripuladas [13]-[15] ou satélites [16], [17].

Neste artigo, utilizou-se sensores embarcados em drones, e busca-se responder a seguinte pergunta: A utilização de um canal infravermelho em conjunto com canais RGB gera um aumento na exatidão global de uma classificação de alvos em imagens geradas por um sensor embarcado em um drone? Este trabalho propõe uma técnica de fusão de dados entre imageamento na banda visível e no infravermelho longo (7,5 a 13,5 micrometros) e verificada a sua exatidão em duas composições diferentes, sejam elas RGB e RGB mais banda

termal (RGBI). A comparação entre as duas composições é feita utilizando os algoritmos k-vizinhos mais próximos (k-NN) e máquina de vetores de suporte (SVM).

O sensor possui a capacidade de capturar simultaneamente imagens e vídeos na banda visível e na banda termal. Entretanto, distorções de imagens ocorrem em virtude da plataforma empregada.

II. MATERIAIS E MÉTODOS

Para aquisição das imagens, foi utilizada a câmera FLIR DUO acoplada ao drone DJI Phantom IV. A câmera possui dois sensores separados, um na faixa visível e outro no infravermelho de ondas longas (LWIR). Por seu tamanho reduzido (41 x 59 x 30 mm) e peso de apenas 84 gramas, seu uso é bastante adequado para a utilização em drones.

Apesar de seu peso reduzido, a câmera não possui alimentação própria, necessitando de uma bateria portátil para sua operação. Embora a bateria utilizada aumente o peso do conjunto em 280 gramas (aproximadamente 20% do peso total do drone) a aquisição das imagens não foi afetada.

O sensor termal é dotado de uma matriz de micro bolômetros sensível na faixa do infravermelho longo, o qual, tem taxa de atualização máxima de 8,33 quadros por segundo no modo filmagem, não necessita de refrigeração e possui ótima resposta para alvos a temperatura ambiente [18], [19]. Demais informações a respeito do sensor podem ser encontradas na Tabela I. A câmera foi acoplada em uma plataforma onde o sensor fica voltado para o nadir. O drone possui um peso de 1380 gramas, com uma autonomia máxima de 28 minutos. O drone possui um recurso de estabilização de imagem que não foi utilizado por incompatibilidade com a plataforma da câmera. As imagens utilizadas foram coletadas em um voo de drone sobre área composta de: edificações, ruas asfaltadas, calçadas, árvores, gramado, carros e pedestres.

TABELA I - ESPECIFICAÇÕES TÉCNICAS DOS SENSORES. FONTE: [20]

-----	<i>Termal</i>	<i>Visível</i>
Faixa de Operação	7,5 a 13,5 μ m	0,38 a 0,74 μ m
Resolução	160 x 120 pixels	1920 x 1080 pixels
Taxa de Atualização máx.	8,33 Hz	30 Hz
FOV	57° x 44°	90°
Níveis Digitais	14 bits	8 bits
Formato da Imagem	.tiff e .jpg	.jpg

A. Metodologia

Ao sobrepor as imagens geradas pelos dois sensores, foi percebido um deslocamento da imagem no visível com relação à imagem termal, conforme demonstrado na Fig. 1. A partir da comparação com diversos outros quadros, somado ao fato de que o conjunto original de imagens na faixa visível possuía duas imagens a mais que as imagens termais, constatou-se que não há sincronia entre a obtenção de imagens nos dois sensores. Um efeito similar foi observado em [21] e [22].

Uma vez que o tempo de integração dos dois sensores é diferente, qualquer vibração da plataforma, seja por vento, correntes térmicas ou mesmo o próprio movimento do drone causará pequenos deslocamentos entre uma banda e outra. Tal efeito é aleatório e o operador não possui qualquer controle sobre ele seja durante o voo assim, para superar esse problema, executou-se uma pré-seleção visual das imagens para remover aquelas que apresentaram esse efeito em maior intensidade. Contudo, uma outra possível solução seria a utilização da técnica de registro de imagens para mitigar esse efeito [23]. Tal aplicação, porém, não foi objetivo deste trabalho.

Ao comparar o retorno infravermelho da superfície metálica do poste de iluminação mostrado na Fig. 1 (região vermelha mais intensa) com a sua posição na imagem na faixa do visível, é possível perceber que na imagem da esquerda o deslocamento se deu para esquerda e para baixo, diferentemente da imagem da direita, na qual o deslocamento foi para a esquerda e para cima.

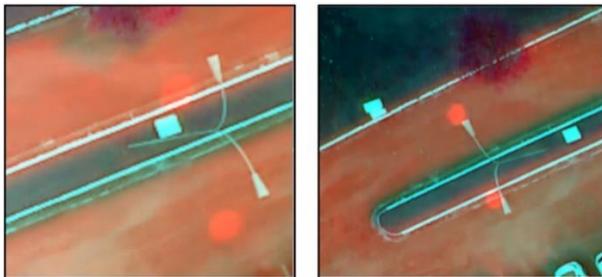


Fig. 1. Comparativo entre dois frames diferentes para demonstração do deslocamento da banda termal com relação à banda visível.

B. Preparação do banco de dados

A imagem do visível foi obtida em formato *widescreen* com um ângulo do campo de visão maior que a imagem do infravermelho. Assim, para sobrepor as imagens das câmeras diferentes foi preciso recortar as bordas laterais da imagem do visível, tal feito é mostrado na Fig. 2 onde pode-se comparar a proporção das imagens. Nela, as linhas tracejadas correspondem ao ponto de corte da imagem, obtendo-se uma largura de 1440 pixels, ou seja, tamanho 9 vezes maior que a imagem do infravermelho. Como as imagens dos dois sensores são significativamente diferentes, optou-se por modificar o tamanho de ambas as imagens para a resolução de 640 x 480 pixels. Nesta resolução, no caso da imagem visível, não há a perda de informações que aconteceria na resolução de 160 x 120 pixels e, por outro lado, não há o custo computacional de se trabalhar com ambas as imagens na resolução de 1440 x 1080 pixels.

Após, foi definido o formato de saída na sobreposição das imagens. Para os testes foi aplicado um *template* com a junção dos canais no tamanho de 640 x 480 pixels. Ressalta-se que, além das transformações do tamanho das imagens, foi necessário converter a imagem do infravermelho para uma resolução de 8 bits. Os canais podem ser observados conforme a Fig. 3. Já na Fig. 4, é possível observar a composição final da imagem com o canal infravermelho substituindo o canal vermelho para permitir a visualização da imagem.

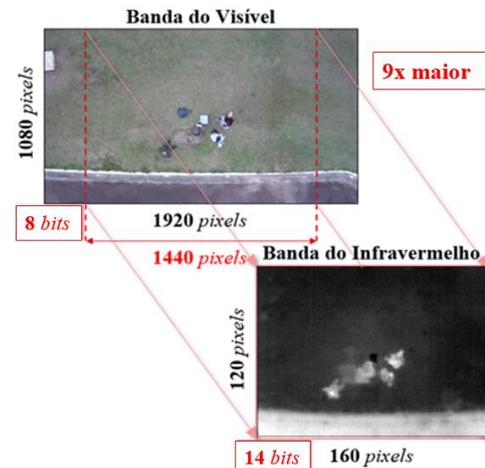


Fig. 2. Comparação da proporção das imagens do visível e do termal.

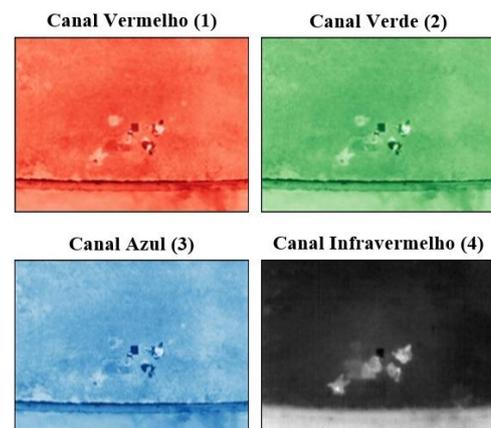


Fig. 3. Composição dos canais das imagens.

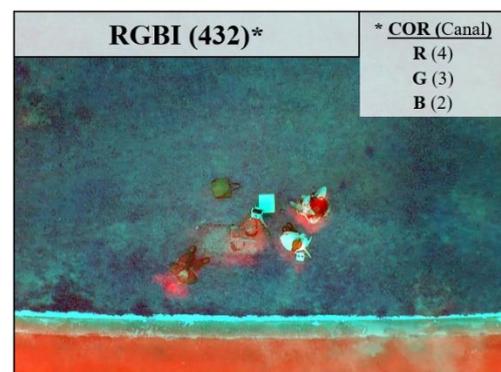


Fig. 4. Composição falsa-cor adotando os canais infravermelho (4), azul (3) e verde (2).

C. *K* Vizinhos mais próximos (*k*-Nearest Neighbors, *k*-NN)

O algoritmo k-NN é uma técnica de classificação supervisionada não paramétrica. O algoritmo não é treinado para produzir um modelo, ele compara cada amostra desconhecida com os dados de treinamento originais. Para isso, requer apenas a configuração do parâmetro k. Ele calcula as distâncias do pixel desconhecido a todos os pixels de treinamento para identificar os k vizinhos mais próximos [24]. Foram testados os valores 3, 5, 7, 9 e 11. O valor que apresentou a melhor exatidão global foi k = 5.

A ideia por trás do método é que os pixels próximos uns dos outros no espaço espectral provavelmente pertencem à mesma classe. Cada amostra desconhecida é atribuída à classe mais comum entre as k amostras de treinamento mais próximas no espaço de atributos.

O k-NN requer um conjunto de treinamento, o qual será utilizado para encontrar as distâncias entre as amostras, por exemplo como a distância euclidiana, dada pela equação:

$$D_E(p, q) = \sqrt{(p_1 - q_1)^2 + \dots + (p_n - q_n)^2} \quad (1)$$

$$D_E(p, q) = \sqrt{\sum_{i=1}^n (p_n - q_n)^2} \quad (2)$$

onde o número de vizinhos mais próximos k será considerado na classificação. A prática mais comum é utilizar um número ímpar para o valor de k, valores muito pequenos são, normalmente, sensíveis à presença de ruídos no conjunto de treinamento. Entretanto, caso o valor seja muito alto, os “vizinhos” tendem a incluir mais elementos de outras classes próximas [24].

D. Máquinas de Vetores de Suporte (Support Vector Machines, SVM)

O algoritmo SVM, assim como o k-NN, é uma técnica de reconhecimento de padrões supervisionada. Consiste em um classificador linear binário não-probabilístico, pois classifica os dados sempre em apenas duas classes. Baseia-se na procura de um hiperplano ótimo, a fim de separar as classes com a maior margem possível.

O treinamento consiste em descobrir a função do hiperplano que separa as classes maximizando a distância da margem que, por sua vez, é determinada pelas amostras de treinamento mais próximas desse hiperplano. Os vetores que partem da origem até essas amostras são chamados *support vectors* [25].

As funções 3 e 4 abaixo definem respectivamente a função do hiperplano e a regra de decisão:

$$g(\vec{x}) = \vec{\omega}^T \vec{x} + b \quad (3)$$

$$\begin{cases} g(\vec{x}) \geq 1, & \forall \vec{x} \in \text{Classe } P \\ g(\vec{x}) \leq -1, & \forall \vec{x} \in \text{Classe } N \end{cases} \quad (4)$$

onde $\vec{\omega}$ é o vetor de pesos, \vec{x} é o vetor de atributos e b é o intercepto [26].

A sua aplicação na classificação de mais de uma classe consiste na utilização de n classificadores binários em que cada

classificador identifica uma classe dentre as classes restantes (*one vs. all*).

E. Seleção de amostras e classes

Dentre as amostras utilizadas no treinamento e teste foram selecionadas quatro classes, entre elas: grama; asfalto, árvore e telhado, está última engloba diferentes tipos de cobertura como telhas de fibrocimento, telhas de metal e concreto. Esperava-se que a combinação das camadas do visível e do infravermelho permitisse uma maior distinção entre elas, já que é aparente a diferença das classes nas imagens segundo a sua tonalidade de cores no RGB e quantidade de energia térmica captada no LWIR. Um exemplo da seleção das amostras de treinamento e teste é mostrado na Fig. 5.

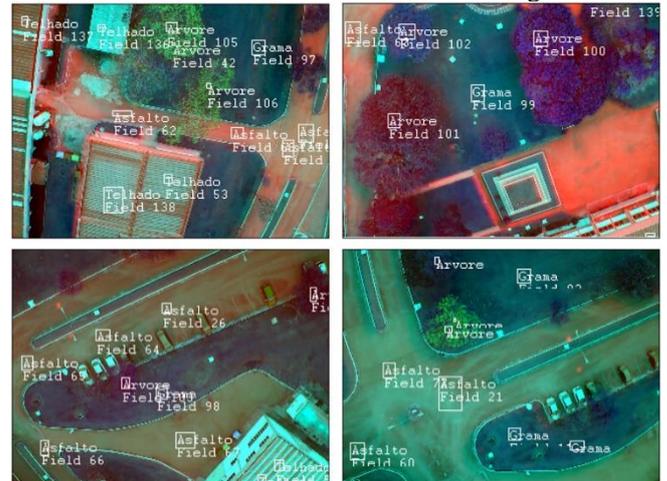


Fig. 5. Seleção de amostras das classes propostas.

III. RESULTADOS E DISCUSSÃO

A Fig. 6 demonstra alguns exemplos dos resultados obtidos na utilização do modelo. É possível perceber que as áreas com asfalto e grama foram classificadas de maneira correta consistentemente. É notável como algoritmo k-NN classificou o meio fio das calçadas e canteiros como pertencentes à classe telhado. Essas classes, apesar de não serem alvo deste estudo, possuíam características semelhantes às das coberturas das construções imageadas.

Pode-se perceber na mesma figura, que o algoritmo SVM confundiu-se mais entre as classes asfalto e grama e também telhado com grama e asfalto. Na análise, isso ocorreu tanto na composição RGB quanto na composição RGBI. As classes asfalto e telhado são semelhantes entre si, justificando os erros de predição entre ambas. Porém, as classes grama e asfalto notadamente possuem diferenças consistente. Um fator que pode ter ocasionado a confusão é a quantidade de detritos por cima do asfalto que também pode ser encontrada na grama em regiões que sua cobertura é mais escassa. Ao adicionar o canal infravermelho, a confusão entre as classes cai para um terço, aproximadamente.

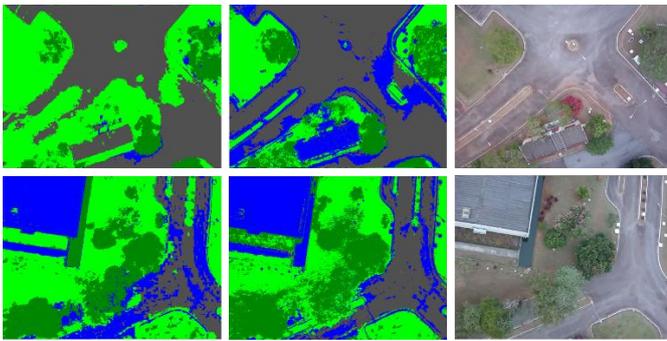


Fig. 6. Comparação das técnicas de processamento. SVM RGBI (à esquerda) e k-NN RGBI (no meio).

Ao analisar a Fig. 7 percebe-se que o modelo k-NN aplicado na composição RGBI apresenta desempenho melhor do que na composição RGB. A classe grama não obteve alterações significativas na exatidão do usuário (94,3% na composição RGBI e 94,9% na composição RGB). A classe asfalto, porém, apresentou uma exatidão do usuário significativamente melhor na composição RGBI com 97,7% contra 92,5% na composição RGB. A classe árvore, assim como a classe grama, não apresentou mudanças relevantes na exatidão do usuário para o modelo k-NN, com 99,2% para a composição RGBI e 98,8% na composição RGB. A classe telhado apresentou um melhor desempenho na exatidão do usuário na composição RGBI (57,2%) contra a composição RGB (36,8%).

De maneira geral é possível perceber que a composição RGBI tem uma melhora significativa na exatidão de usuário e produtor para as classes asfalto e telhado, consequentemente, a exatidão global da composição RGBI ficou em 89,7% contra 80,0% da composição RGB. A estatística kappa, também, comprova a melhor exatidão entre as classes do modelo para a composição RGBI.

A Fig. 8 apresenta os resultados relativos ao algoritmo SVM. Com desempenho significativamente inferior ao algoritmo k-NN para o mesmo conjunto de dados, observa-se que a composição RGBI não obteve vantagem significativa da acurácia em relação à composição RGB (82,5% contra 80,6% respectivamente).

Comparando-se as duas composições, a classe grama teve uma exatidão de usuário maior na composição RGBI (87%) com relação à composição RGB (78,7%). A classe asfalto apresentou exatidão do usuário de 91,4% para a composição RGBI contra 87,5% para a composição RGB. As classes árvore e telhado não apresentaram alterações significativas de exatidão de usuários na composição RGB para RGBI.

Ainda na Fig. 8, é possível notar que a classe telhado não se confunde com a classe árvore. Mesmo sem a informação do canal infravermelho, essas classes apresentaram uma boa separabilidade por apresentarem respostas espectrais diferentes na banda do visível e do infravermelho longo.

Percebe-se também, que a exatidão do produtor para a classe telhado apresentou um aumento significativo partindo de 58,9% na composição RGB para 73% na composição RGBI, muito embora essa classe tenha apresentado o pior desempenho em ambos os modelos.

k-NN (RGBI)

		PREDIÇÃO				Amostras	EU (%)
		Gramma	Asfalto	Arvore	Telhado		
Gramma	6636	8	392	0	7036	94.3	
Asfalto	0	7208	0	170	7378	97.7	
Árvore	28	0	3549	0	3577	99.2	
Telhado	0	1648	0	2200	3848	57.2	
TOTAL	6664	8864	3941	2370	21839		
EP (%)	99.6	81.3	90.1	92.8			
Acurácia	89.7%	Estatística Kappa			85.6%		

k-NN (RGB)

		PREDIÇÃO				Amostras	EU (%)
		Gramma	Asfalto	Arvore	Telhado		
Gramma	6496	16	336	0	6848	94.9	
Asfalto	126	5303	0	304	5733	92.5	
Árvore	42	0	3605	0	3647	98.8	
Telhado	0	3545	0	2066	5611	36.8	
TOTAL	6664	8864	3941	2370	21839		
EP (%)	97.5	59.8	91.5	87.2			
Acurácia	80.0%	Estatística Kappa			73.0%		

Fig. 7. Matriz de confusão do modelo k-NN nas composições RGBI e RGB.

SVM (RGBI)

		PREDIÇÃO				Amostras	EU (%)
		Gramma	Asfalto	Arvore	Telhado		
Gramma	6104	424	490	0	7018	87.0	
Asfalto	0	6736	0	635	7371	91.4	
Árvore	133	0	3451	5	3589	96.2	
Telhado	427	1704	0	1730	3861	44.8	
TOTAL	6664	8864	3941	2370	21839		
EP (%)	91.6	76.0	87.6	73.0			
Acurácia	82.5%	Estatística Kappa			75.6%		

SVM (RGB)

		PREDIÇÃO				Amostras	EU (%)
		Gramma	Asfalto	Arvore	Telhado		
Gramma	6034	1144	490	0	7668	78.7	
Asfalto	0	6728	0	965	7693	87.5	
Árvore	42	0	3451	10	3503	98.5	
Telhado	588	992	0	1395	2975	46.9	
TOTAL	6664	8864	3941	2370	21839		
EP (%)	90.5	75.9	87.6	58.9			
Acurácia	80.6%	Estatística Kappa			72.6%		

Fig. 8. Matriz de confusão do modelo SVM nas composições RGBI e RGB.

A exatidão global da composição RGBI do algoritmo SVM foi de 82,5% e na composição RGB foi de 80,6 %, em suma, não se obteve melhora significativa no método empregado. A estatística kappa entre as composições também comprova tal fato.

Do modo geral, percebeu-se uma melhora da exatidão global com o acréscimo do canal infravermelho aos canais RGB. Apesar disso, com uma análise mais profunda em algumas das imagens, percebe-se que o erro do deslocamento de entre as sobreposições de camadas pode prejudicar pontos específicos da imagem, como canteiros centrais de algumas ruas, conforme mostra a Fig. 9.

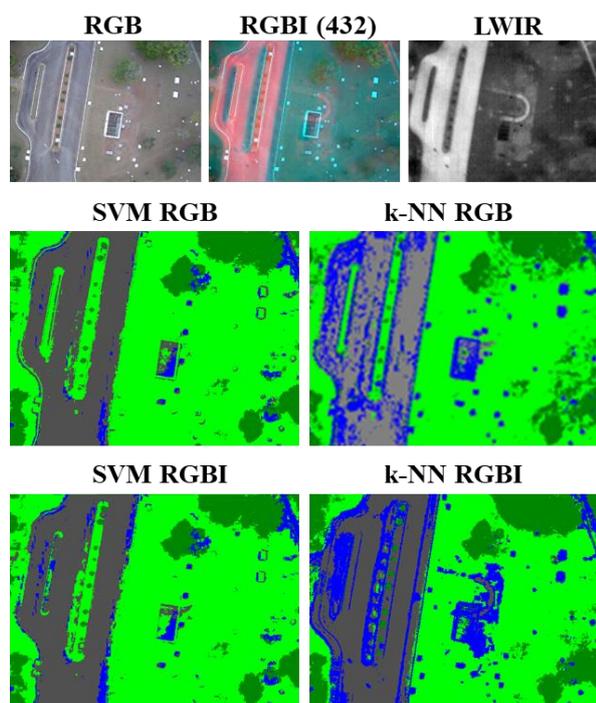


Fig. 9. Comparação das composições de um frame em RGB, RGBI (camadas 4, 3 e 2) e infravermelho com a respectiva classificação após a predição nos modelos k-NN e SVM em ambas as composições

A classificação do modelo da camada do visível tem maior exatidão de usuário na classe grama e árvore, entretanto, nas classes telhado e asfalto, o desempenho desse parâmetro foi inferior ao da composição RGBI. Percebe-se que, apesar do deslocamento entre as camadas do visível e termal, como as classes telhado e asfalto apresentam maiores intensidades da camada termal, a composição RGBI consegue diferenciar melhor essas duas classes. De maneira inversa, como as classes grama e árvore apresentam intensidades mais próximas na camada infravermelho, a adição desse canal não proporciona melhora na classificação.

De fato, a inclusão da camada do infravermelho no modelo permitiu uma maior exatidão, tanto de usuário como de produtor, para algumas das classes analisadas. Contudo, o deslocamento aleatório entre a imagem infravermelho e a imagem RGB pode causar um erro na classificação nas áreas afetadas da imagem. Vale ressaltar que não foi o objetivo deste trabalho a correção da sobreposição entre as camadas.

IV. CONCLUSÕES

Este trabalho buscou responder se a utilização em Sensoriamento Remoto de um canal infravermelho em conjunto com canais RGB gera um aumento na exatidão global de uma classificação de alvos em imagens aéreas. Tendo como base os resultados apresentados e discutidos, pode-se perceber que a utilização de um quarto canal com dados do sensor infravermelho permitiu um melhor desempenho, com relação à exatidão global, na identificação das classes estudadas. Ressalta-se que o método empregado não é exclusivo para uso em drones, mas de qualquer meio que combine um quarto canal no processo de captura de imagens.

Pode-se concluir que é viável e vantajosa a fusão de sensores para classificação de imagens pois, mesmo com o deslocamento do canal infravermelho com relação aos canais RGB, as imagens na composição RGBI apresentaram uma melhor exatidão global na aplicação dos algoritmos k-NN e SVM apesar de que, neste último, a melhora não foi significativa. Pode-se inferir também que aplicação de outra camada de informação aumenta a consciência situacional sobre o cenário de operação, o que gera uma vantagem significativa sobretudo em aplicações na área de defesa.

É possível perceber o melhor desempenho na identificação das classes que possuem intensidade significativa no infravermelho longo (asfalto e telhado) e uma redução nas classes que não possuem diferenciação adequada nesse mesmo canal (grama e árvore).

Uma evolução do trabalho apresentado seria o estudo sobre a aplicação de algoritmos para a correção do deslocamento entre os canais das imagens que poderia proporcionar uma melhora nos resultados. Assim, seria possível avaliar a diferença de desempenho dos algoritmos de classificação em comparação às imagens sem correção.

v. REFERÊNCIAS

- [1] Mohd Noor, N., Abdullah, A. and Hashim, M., 2018. Remote sensing UAV/drones and its applications for urban areas: A review, IOP Conference Series: Earth and Environmental Science 2018.
- [2] Ludovisi, R., Tauro, F., Salvati, R., Khoury, S., Mugnozza, G.S. and Harfouche, A., 2017. Uav-based thermal imaging for high-throughput field phenotyping of black poplar response to drought. *Frontiers in Plant Science*, 8.
- [3] Henry, C., Poudel, S., Lee, S.-. and Jeong, H., 2020. Automatic detection system of deteriorated PV modules using drone with thermal camera. *Applied Sciences (Switzerland)*, 10(11).
- [4] Sambolek, S. and Ivacic-Kos, M., 2021. Automatic person detection in search and rescue operations using deep CNN detectors. *IEEE Access*, 9, pp. 37905-37922.
- [5] Rodin, C.D., De Lima, L.N., De Alcantara Andrade, F.A., Haddad, D.B., Johansen, T.A. and Storvold, R., 2018. Object Classification in Thermal Images using Convolutional Neural Networks for Search and Rescue Missions with Unmanned Aerial Systems, *Proceedings of the International Joint Conference on Neural Networks 2018*.
- [6] Andrade, F. A. de A. et al. Autonomous unmanned aerial vehicles in search and rescue missions using real-time cooperative model predictive control. *Sensors (Switzerland)*, v. 19, n. 19, 2019. DOI 10.3390/s19194067.
- [7] Rasmussen, N. D. et al. "Fused visible and infrared video for use in wilderness search and rescue." 2009. *Anais [...]. [s.l.: s.n.]*, 2009. DOI 10.1109/WACV.2009.5403048.
- [8] 171. St-Laurent, L., Maldague, X., Prevost, D.: "Combination of colour and thermal sensors for enhanced object detection." In: 10th International Conference on Information Fusion (2007).
- [9] Klein, L.A. 2012, "Sensor and data fusion: A tool for information assessment and decision making: Second edition" in *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making: Second Edition*, pp. 1-474.
- [10] A. J. A. Rivera, A. D. C. Villalobos, J. C. N. Monje, J. A. G. Mariñas, and C. M. Oppus, "Post-disaster rescue facility: Human detection and geolocation using aerial drones," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, pp. 384-386, Feb. 2017, doi: 10.1109/TENCON.2016.7848026.
- [11] H. Kayan, R. Eslampanah, F. Yeganli, and M. Askar, "Heat leakage detection and surveillance using aerial thermography drone," *26th IEEE Signal Process. Commun. Appl. Conf. SIU 2018*, pp. 1-4, Jul. 2018, doi: 10.1109/SIU.2018.8404366.
- [12] P. M. Hell and P. J. Varga, "Assisting law enforcement tasks with thermal camera drones," *CANDO-EPE 2020 - Proceedings, IEEE 3rd Int. Conf. Work. Obuda Electr. Power Eng.*, pp. 97-102, Nov. 2020, doi: 10.1109/CANDO-EPE51100.2020.9337768.

- [13] H. Greidanus, "Assessment of the coastal maritime environment with airborne mid-wave infrared imagery," *Int. Geosci. Remote Sens. Symp.*, vol. 3, pp. 1542–1544, 2001, doi: 10.1109/IGARSS.2001.976905.
- [14] M. Sadi, Y. Zhang, W. F. Xie, and F. M. A. Hossain, "Forest Fire Detection and Localization Using Thermal and Visual Cameras," *2021 Int. Conf. Unmanned Aircr. Syst. ICUAS 2021*, pp. 744–749, Jun. 2021, doi: 10.1109/ICUAS51884.2021.9476865.
- [15] J. M. Suiter, M. B. Lapis, and R. Collins, "Multispectral image fusion for the small aircraft transportation system," *AIAA/IEEE Digit. Avion. Syst. Conf. - Proc.*, vol. 1, 2004, doi: 10.1109/DASC.2004.1391317.
- [16] R. Hartono, D. Ardianto, S. Salaswati, R. Yatim, and A. Hadi Syafrudin, "Design Requirement of LWIR Optical Filter for LAPAN-A4 Satellite," *Proc. 2019 IEEE Int. Conf. Aerosp. Electron. Remote Sens. Technol. ICARES 2019*, Oct. 2019, doi: 10.1109/ICARES.2019.8914354.
- [17] B. Arifin, A. M. Tahir, and I. Priyanto, "LAPAN'S Mid Wavelength Infrared Camera Module," *Int. Geosci. Remote Sens. Symp.*, pp. 6401–6404, Sep. 2020, doi: 10.1109/IGARSS39084.2020.9323882.
- [18] Gade, R.; Moeslund, T. B. *Thermal cameras and applications: A survey. Machine Vision and Applications*, v. 25, n. 1, p. 245–262, 2014. DOI 10.1007/s00138-013-0570-5.
- [19] Serway, R.A., Jewett, J.W.: *Physics for Scientists and Engineers with Modern Physics*, 6th edn. Brooks/Cole–Thomson Learning (2004).
- [20] Duo-Datasheet-US. Disponível em: <<https://flir.netx.net/file/asset/10904/original/attachment>>. Acesso em: 05/07/2021.
- [21] Irani, M.; Anandan, P. "Robust Multi-Sensor Image Alignment." *Sixth International Conference on Computer Vision*, 1998.
- [22] Istenic, R. et al. "Thermal and visual image registration in hough parameter space." 2007. p. 106–109. DOI 10.1109/IWSSIP.2007.4381164.
- [23] Zitová, B. & Flusser, J. 2003, "Image registration methods: A survey", *Image and Vision Computing*, vol. 21, no. 11, pp. 977-1000. DOI 10.1016/S0262-8856(03)00137-9.
- [24] Richards, John A.; JIA, Xiuping. "Remote Sensing Digital Image Analysis: An Introduction." 4. ed. New York: Springer, 2005. 464 p. ISBN 9783540251286.
- [25] Falqueto, L. E. F., "Reconhecimento de alvos artificiais em ambientes complexos por meio de imagens SAR polarimétricas obtidas por plataformas orbitais, 2019, 133f. Dissertação de mestrado em Ciências e Tecnologias Espaciais – Instituto Tecnológico de Aeronáutica, São José dos Campos.
- [26] Kotsiantis, S. B; Zaharakis, I. D; Pintelas, P. E., "Machine Learning: a review of classification and combining techniques. *Artificial Intelligence Review*, v. 26, n. 3, p. 159-190, 2006.