UDC 004.932

CORRELATION-BASED FUSION OF FEATURES FOR SATELLITE IMAGE OBJECT CLASSIFICATION

[1*]*Fazilov Sh.Kh.*, [2]*Yusupov O.R.*, [1]*Khandamov Y.Kh.*

[*]sh.fazilov@gmail.com
[1]Digital technologies and artificial intelligence development research institute,
Tashkent, Uzbekistan.
[2]Samarkand State University named after Sharof Rashidov,
Samarkand, Uzbekistan

Abstract. This study analyzes feature fusion methods based on Convolutional Neural Networks (CNN) and Canonical Correlation Analysis to enhance the effectiveness of vehicle detection and classification in remote sensing images. The experimental investigations were conducted on the Cars Overhead With Context (COWC) dataset. Initially, fine-tuning was performed on ResNet18, DenseNet121, and VGG16 models using 20% of the data, and feature vectors were extracted from their output layers. Each model was evaluated individually, as well as through feature fusion using various statistical approaches–CCA (Canonical Correlation Analysis), DCCA (Discriminative Canonical Correlation Analysis), MCCA (Multiple Canonical Correlation Analysis), and DMCCA (Discriminative Multiple Canonical Correlation Analysis). The results demonstrated that among individual CNN models, DenseNet121 achieved the highest accuracy (≈87.45%). However, the features fused via DMCCA, when classified with an SVM, achieved the best overall performance with an accuracy of approximately 91.52%, along with superior precision, recall, and F1-score values.

Keywords: Remote sensing, COWC, Discriminative Multiple Canonical Correlation Analysis, DMCCA, Feature Fusion, SVM, Classification.

1 Introduction

The rapid growth of data volume and the diversity of information coming from various sources today demand new approaches in scientific research and practical systems. For example, in remote sensing, images are captured across different spectral ranges; in biomedical fields, multiple diagnostic signals such as EEG, MRI, and ECG are utilized; while in multimedia systems, image, speech, and text modalities are often combined simultaneously. In such contexts, feature fusion methods play a crucial role [1]. The main objective of feature fusion is to represent features obtained from different sources or models within a unified space, thereby exploiting more information, enhancing discriminative power, and improving classification accuracy. Statistical methods are among the most reliable and theoretically grounded approaches for addressing this task [2]. This paper provides a comprehensive overview of statistical methods applied in feature fusion, analyzing their theoretical foundations, practical implementations, advantages, and limitations.

Feature fusion is the process of combining features extracted from multiple data sources or models into a single feature vector. This process can be performed at three levels: Early Fusion, which directly merges raw data or low-level features; Intermediate Fusion, which projects features into a new space through various transformations or statistical mappings; and Late Fusion, which statistically combines decisions at the classifier output level [3].

Statistical approaches are widely used at the intermediate fusion stage because they transform feature spaces into more compact, discriminative, and noise-robust representations. Among the commonly applied statistical methods are Principal Component Analysis (PCA) [4], Linear/Fisher Discriminant Analysis (LDA/FDA) [5], Canonical Correlation Analysis (CCA) [6], Discriminative Canonical Correlation Analysis (DCCA) [7], Multiple Canonical Correlation Analysis (MCCA) [8], and Discriminative Multiple Canonical Correlation Analysis (DMCCA) [9].

Problem Statement

Let us assume that we have a set of features extracted from different sources (modalities) or multiple models:

$$X^{(1)} \in R^{n\times m_1}, X^{(2)} \in R^{n\times m_2}, \dots, X^{(p)} \in R^{n\times m_P}, \quad (1)$$

where n denotes the number of objects, m_P is the dimensionality of the features obtained from the p-th source, and $X^{(p)} = [x_1^{(p)}, x_2^{(p)}, \dots, x_n^{(p)}]^T$ represents the feature matrix corresponding to the the p-th source.

The general objective of feature fusion is to project these features into a unified latent space and thereby enhance the ability to distinguish (classify) different classes. For this purpose, an appropriate projection matrix $W^{(p)} \in R^{m_P\times d}$ is determined for each source, and the projected features are obtained as

$$Y^{(p)} = W^{(p)T}X^{(p)}, p = 1,2, \dots, P. \quad (2)$$

Based on these projections, a common fused feature representation is formed as

$$Z = f\left(Y^{(1)}, Y^{(2)}, \dots, Y^{(p)}\right) \in R^{n\times d}, \quad (3)$$

where $f(\cdot)$ is a fusion function (e.g., concatenation, summation, or weighted combination), and d denotes the dimensionality of the common latent space ($d \ll m_P$).

In general, the feature fusion problem can be expressed as the following optimization task:

$$\max_{W^{(1)},\dots,W^{(p)}} \mathcal{J}(f(Y^{(1)}), f(Y^{(2)}), \dots, f(Y^{(p)})) \quad (4)$$

where $\mathcal{J}(\cdot)$ represents the selected statistical criterion–such as variance correlation (CCA, MCCA), or discriminative correlation (DCCA, DMCCA)).

2 Solution methods

The feature fusion problem aims to represent information obtained from multiple sources within a common latent space and to improve classification performance. The solution methods can generally be divided into two categories:

- Fusion of two sources of features (CCA, DCCA)
- Fusion of multiple sources of features (MCCA, DMCCA).

3.1 Fusion of Two Sources of Features (CCA, DCCA)

In many data processing scenarios, features obtained from multiple sources (e.g., sensors or different data formats) are available. For instance, optical and radar images of the same region, or textual descriptions and corresponding images, may represent different modalities of information. Although these data sources differ in form, fusing their features can significantly enhance the overall quality of classification or analysis. Therefore, two-source feature fusion (also referred to as *bimodal feature fusion*) represents an essential component of multimodal analysis.

3.1.1 Feature Fusion Based on CCA

Canonical Correlation Analysis (CCA) is a statistical method used to identify the maximal correlations between two different sources of data. First introduced by H. Hotelling in 1936, its main idea is to project both data sources into a new space through suitable linear transformations, such that the correlation between the projected features is maximized. Thus, CCA reveals the underlying common structure between two data sources and enables the identification of their strongest interrelationships.

Let two sets of features from different sources be given as:

$$X^{(1)} \in R^{n\times m_1}, X^{(2)} \in R^{n\times m_2} \quad (6)$$

The optimization problem of CCA can be formulated as:

$$\max_{W^{(1)},W^{(2)}} \mathcal{J}\left(W^{(1)}, W^{(2)}\right) = \max_{W^{(1)},W^{(2)}} \frac{{W^{(1)}}^T C_{12} W^{(1)}}{\sqrt{{W^{(1)}}^T C_{11} W^{(1)}} \cdot \sqrt{{W^{(2)}}^T C_{22} W^{(2)}}} \quad (7)$$

where

$$C_{11} = \frac{1}{n} X^{(1)^T} X^{(1)}, C_{22} = \frac{1}{n} X^{(2)^T} X^{(2)}, C_{12} = \frac{1}{n} X^{(1)^T} X^{(2)}, C_{21} = \frac{1}{n} X^{(2)^T} X^{(1)} \quad (8)$$

Maximizing the above ratio leads to the following pair of generalized eigenvalue problems:

$$\begin{aligned} C_{12} C_{22}^{-1} C_{21} W^{(1)} &= \lambda^2 C_{11} W^{(1)}, \\ C_{21} C_{11}^{-1} C_{12} W^{(2)} &= \lambda^2 C_{22} W^{(2)}. \end{aligned} \quad (9)$$

Here, $W^{(1)}$ and $W^{(2)}$ denote the projection matrices for the two sources. Using these matrices, the projected features are obtained as:

$$\begin{aligned} Z^{(1)} &= W^{(1)^T} X^{(1)}, \\ Z^{(2)} &= W^{(2)^T} X^{(2)}. \end{aligned} \quad (10)$$

Finally, the fused feature vector is constructed as

$$Y = f\left(Z^{(1)}, Z^{(2)}\right) \in R^{n \times d} \quad (11)$$

where f denotes the feature fusion function.

3.1.2 Feature Fusion Based on DCCA

Discriminative Canonical Correlation Analysis (DCCA) is a discriminative extension of CCA that enables more effective integration of multimodal data according to class information. It maximizes within-class correlations while minimizing between-class correlations, thereby achieving higher performance in classification and identification tasks.

Let two sets of features be given:

$$X^{(1)} \in R^{n \times m_1}, X^{(2)} \in R^{n \times m_2} \quad (12)$$

and the corresponding class labels be

$$y_i \in \{1, \dots, c\} \quad (13)$$

The DCCA optimization problem can be formulated as follows:

$$\max_{W^{(1)}, W^{(2)}} \mathcal{J}\left(W^{(1)}, W^{(2)}\right) =$$

$$\max_{W^{(1)}, W^{(2)}} \left(\frac{W^{(1)^T} C_{12}^{(w)} W^{(2)}}{\sqrt{W^{(1)^T} C_{11}^{(w)} W^{(1)}} \cdot \sqrt{W^{(2)^T} C_{22}^{(w)} W^{(2)}}} - \beta \frac{W^{(1)^T} C_{12}^{(b)} W^{(2)}}{\sqrt{W^{(1)^T} C_{11}^{(b)} W^{(1)}} \cdot \sqrt{W^{(2)^T} C_{22}^{(b)} W^{(2)}}} \right) \quad (14)$$

where $C^{(w)}$ and $C^{(b)}$ denote the within-class and between-class covariance matrices, respectively, and β is a balancing parameter that controls the trade-off between within-class and between-class correlations.

To solve this optimization problem, the following block generalized eigenvalue problem is formulated:

$$\begin{bmatrix} 0 & C_{12} \\ C_{21} & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \rho \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad (15)$$

As a result, the projection matrices $W^{(1)}$ and $W^{(2)}$ are obtained, and the projected features are defined as:

$$\begin{aligned} Z^{(1)} &= W^{(1)^T} X^{(1)}, \\ Z^{(2)} &= W^{(2)^T} X^{(2)}. \end{aligned} \quad (16)$$

Finally, the fused feature vector is constructed as:

$$Y = f\left(Z^{(1)}, Z^{(2)}\right) \in R^{n \times d} \quad (17)$$

where $f(\cdot)$ denotes the feature fusion function.

3.1.2 Fusion of Multiple Sources of Features (MCCA, DMCCA)

Multi-source feature fusion is one of the key directions in modern data analysis. By integrating features obtained from different sources, it becomes possible to achieve richer, more stable, and more discriminative representations. The choice of the fusion method depends on several factors, including the number of sources, the availability of labels, computational resources, and the nature of the task.

The primary goal of this approach is to represent features from multiple sources within a unified space and to generate a more informative and discriminative feature vector for subsequent analysis tasks such as classification, clustering, and identification.

3.2.1 Feature Fusion Based on MCCA

Multiple Canonical Correlation Analysis (MCCA) is a generalized extension of the classical CCA method designed to handle three or more data sources. While CCA aims to maximize the correlation between two sources, MCCA simultaneously considers the overall correlation among all available sources. It has been successfully applied in various domains such as remote sensing, where it enables the fusion of multispectral and hyperspectral images; biometrics, through the integration of features such as face, voice, and fingerprint; and multimodal sentiment analysis, by combining textual, audio, and video features. Through this comprehensive approach, MCCA provides a more complete representation of shared information across multiple modalities, thereby enhancing robustness and generalization in downstream applications.

Let us assume that we have a set of features obtained from different sources or multiple models:

$$X^{(1)} \in R^{n\times m_1}, X^{(2)} \in R^{n\times m_2}, \dots, X^{(p)} \in R^{n\times m_P} \tag{18}$$

First, both the within-source and between-source covariance matrices are computed as follows:

$$\begin{aligned} C_{pp} &= \frac{1}{n} {X^{(p)}}^T X^{(p)}, p = 1, \dots, P; \\ C_{pq} &= \frac{1}{n} {X^{(p)}}^T X^{(q)}, p \neq q. \end{aligned} \tag{19}$$

Hence, the overall block covariance matrix can be expressed in the following form:

$$C = \begin{bmatrix} C_{11} C_{12} \cdots C_{1P} \\ C_{21} C_{22} \cdots C_{2P} \\ \vdots \quad \vdots \quad \ddots \quad \vdots \\ C_{P1} C_{P2} \cdots C_{PP} \end{bmatrix}. \tag{20}$$

The MCCA optimization problem is formulated as:

$$\max_{W^{(1)},\dots,W^{(P)}} \mathcal{J}\left(W^{(1)}, \dots, W^{(P)}\right) = \max_{W^{(1)},\dots,W^{(P)}} \frac{\sum_{p\neq q} {W^{(p)}}^T C_{pq} W^{(q)}}{\sum_{p=1}^{P} {W^{(p)}}^T C_{pp} W^{(p)}}. \tag{21}$$

This optimization leads to the following generalized block eigenvalue problem:

$$(C - D)W = \lambda DW, \tag{22}$$

where C is the total covariance matrix and D is the block-diagonal matrix that contains only the within-source covariance terms. The solution is obtained by selecting the eigenvectors corresponding to the largest eigenvalues, which define the projection matrices $W^{(p)}$.

The projected features are then computed as

$$Z^{(p)} = {W^{(p)}}^T X^{(p)}, p = 1, \dots, P \tag{23}$$

and the fused feature representation is obtained as

$$Y = f\left(Z^{(1)}, Z^{(2)}, \dots, Z^{(p)}\right) \in R^{n\times d} \tag{24}$$

3.2.2 Feature Fusion Based on DMCCA

Discriminative Multiple Canonical Correlation Analysis (DMCCA) is a discriminative extension of the MCCA method that incorporates class information into the feature fusion process. While MCCA maximizes the overall correlation among multiple sources, DMCCA additionally enhances the discriminative power of the projections by considering class-specific information.

Let us assume that we have a set of features obtained from different sources or multiple models:

$$X^{(1)} \in R^{n\times m_1}, X^{(2)} \in R^{n\times m_2}, \dots, X^{(p)} \in R^{n\times m_P}. \tag{25}$$

First, both the within-source and between-source covariance matrices are computed as follows:

$$C_{pp} = \frac{1}{n} X^{(p)^T} X^{(p)}, p = 1, \dots, P;$$
$$C_{pq} = \frac{1}{n} X^{(p)^T} X^{(q)}, p \neq q. \tag{26}$$

The overall block covariance matrix can then be written as:

$$C = \begin{bmatrix} C_{11} C_{12} \cdots C_{1P} \\ C_{21} C_{22} \cdots C_{2P} \\ \vdots \quad \vdots \quad \ddots \quad \vdots \\ C_{P1} C_{P2} \cdots C_{PP} \end{bmatrix}. \tag{27}$$

Next, for each source p, the within-class and between-class covariance matrices are defined as:

$$S_w^{(p)} = \sum_{j=1}^{C} \sum_{x_i \in c_j} \left(X_i^{(p)} - \mu_j^{(p)} \right) \left(X_i^{(p)} - \mu_j^{(p)} \right)^T,$$
$$S_b^{(p)} = \sum_{j=1}^{C} n_j \left(\mu_j^{(p)} - \mu^{(p)} \right) \left(\mu_j^{(p)} - \mu^{(p)} \right)^T, \tag{28}$$

where $X_i^{(p)}$ is the feature vector of the i-th sample from the p-th source, $\mu_j^{(p)}$ is the class mean of class j for the p-th source, and $\mu^{(p)}$ is the global mean vector.

The block-structured within-class and between-class covariance matrices are expressed as:

$$S_w = \begin{bmatrix} S_w^{(1)} & 0 & \cdots & 0 \\ 0 & S_w^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_w^{(P)} \end{bmatrix}; \tag{29}$$

$$S_b = \begin{bmatrix} S_b^{(1)} & 0 & \cdots & 0 \\ 0 & S_b^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_b^{(P)} \end{bmatrix}. \tag{30}$$

The DMCCA optimization problem is formulated as:

$$\max_{W^{(1)}, \dots, W^{(P)}} \mathcal{J}\left(W^{(1)}, \dots, W^{(P)}\right) = \max_{W^{(1)}, \dots, W^{(P)}} \frac{\sum_{p \neq q} W^{(p)^T} C_{pq} W^{(q)} + \alpha W^T S_b W}{\sum_{p=1}^{P} W^{(p)^T} C_{pp} W^{(p)} + \alpha W^T S_w W} \tag{31}$$

where α is a regularization parameter that controls the trade-off between overall correlation and discriminative ability.

This problem can be transformed into the following generalized eigenvalue formulation:

$$(C - D + \alpha S_b) W = \lambda (D + \alpha S_w) W \tag{32}$$

where C is the total covariance matrix, and D is a block-diagonal matrix that contains only the within-source covariance blocks.

The solution is obtained by selecting the eigenvectors corresponding to the largest eigenvalues, which define the projection matrices $W^{(p)}$. The projected feature sets are then computed as

$$Z^{(p)} = W^{(p)^T} X^{(p)}, p = 1, \dots, P \tag{33}$$

and the fused feature representation is formed as

$$Y = f\left(Z^{(1)}, Z^{(2)}, \dots, Z^{(p)}\right) \in R^{n \times d} \tag{34}$$

3 Results

All experiments were conducted on a computer with the following technical specifications: Windows 10 (64-bit), Intel(R) Core(TM) i9-12900K CPU @ 5.2 GHz, 64 GB RAM, and an NVIDIA RTX 3090 Ti GPU with 24 GB of memory.

To evaluate the effectiveness of the aforementioned feature fusion approaches, the Cars Overhead With Context (COWC) [10] dataset was utilized. This publicly available dataset is designed for vehicle detection from aerial imagery. The dataset consists of images collected from six different sources, with a spatial

resolution of approximately 15 cm per pixel. The vehicles in the dataset range in size from 24 to 48 pixels.The objects in the dataset are categorized into four classes: car, pickup, uncertain vehicle type, and car-like object.



Fig. 1. Samples from the COWC dataset

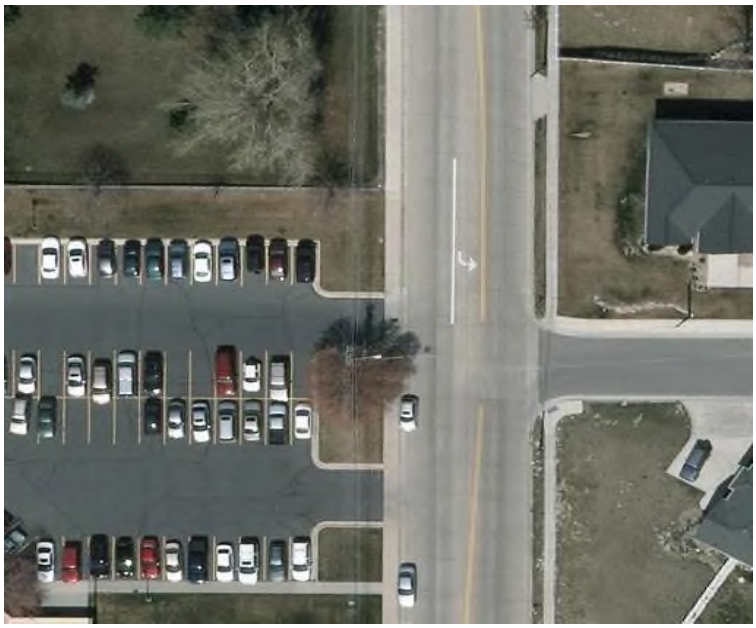


Fig. 2. Example of a remote sensing image from the COWC dataset used for vehicle identification

In the experiments, several convolutional neural network (CNN) models – ResNet18, DenseNet121, and VGG16 – were employed along with different feature fusion methods. The dataset was divided into training and testing subsets with an 20/80 split. Each CNN model was individually fine-tuned on 20% of the data, and the features were extracted from the fully connected layer of each network. These extracted features were then fused using various correlation-based fusion methods, namely CCA, DCCA, MCCA, and DMCCA. The fused feature representations were used for classification with a Support Vector Machine (SVM). For performance evaluation, the following metrics were used: Accuracy, Precision, Recall, and F1-score. All models were tested on the same set

of images, ensuring a fair and consistent comparison. The experimental results are presented in Table 1.

Tab. 1. Classification results on the COWC dataset

№	Model	Feature Fusion	Accuracy	Precision	Recall	F1-score
1.	ResNet18	Individual	86.12	85.9	86	85.95
2.	DenseNet121	Individual	87.45	87.2	87.3	87.25
3.	VGG16	Individual	85.76	85.6	85.7	85.65
4.	ResNet18+ DenseNet121	CCA	88.54	88.3	88.4	88.35
5.	ResNet18+ DenseNet121	DCCA	89.63	89.4	89.5	89.45
6.	ResNet18+ DenseNet121 + VGG16	MCCA	90.21	90	90.1	90.05
7.	ResNet18+ DenseNet121 + VGG16	DMCCA	91.52	91.09	91.21	91.14

As shown in the table, although the individual CNN models produced relatively close results, feature fusion significantly improved overall performance. The lowest accuracy was obtained by VGG16 (85.76%), while DenseNet121 achieved the highest individual accuracy of 87.45%. When the features of ResNet18 and DenseNet121 were fused using CCA, the accuracy increased to 88.54%, indicating that identifying correlation between two feature sources via CCA plays an important role in enhancing classification performance. However, the DCCA approach further improved the results, reaching an accuracy of 89.63%. This demonstrates that DCCA, unlike CCA, additionally considers inter-class discriminative information. When combining features from all three models using MCCA, the accuracy rose to 90.21%, confirming the benefit of utilizing multi-source correlation. Finally, the DMCCA approach achieved the best overall performance with 91.52% accuracy and an F1-score of 91.35%. These results clearly highlight the main advantage of DMCCA: it not only maximizes inter-source correlations across multiple modalities but also enhances their class separability (discriminative capability). The effectiveness of DMCCA-based feature fusion can also be observed in the confusion matrix shown in Figure 3.

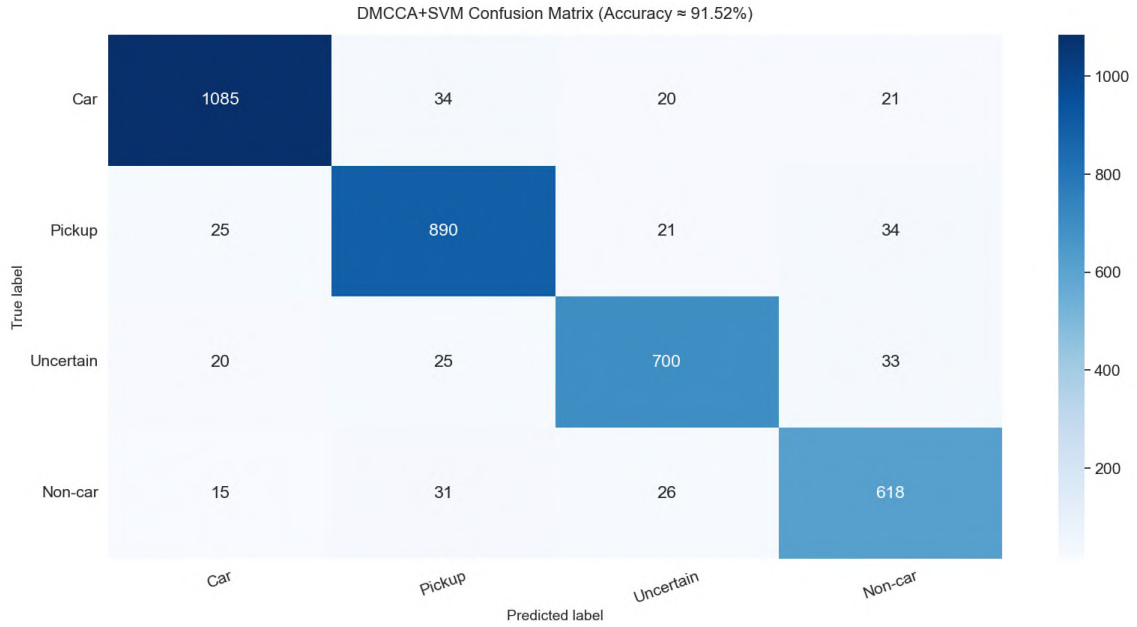


Fig. 3. Confusion matrix for the COWC dataset using the DMCCA + SVM method.

4 Conclusion

In this study, the effectiveness of deep learning–based models and their discriminative correlation fusion methods for remote sensing image analysis was comprehensively investigated using the Cars Overhead With Context (COWC) dataset. During the experiments, features extracted from the ResNet18, DenseNet121, and VGG16 architectures were evaluated both individually and through various fusion approaches, including CCA, DCCA, MCCA, and DMCCA. The results demonstrated that although each individual CNN model achieved high accuracy (around 87%), combining their features by considering inter-model correlations significantly improved performance.

In particular, the DMCCA approach achieved the highest results on the COWC dataset – Accuracy: 91.52%, Precision: 91.09%, Recall: 91.21%, and F1-score: 91.14%. These findings indicate that DMCCA not only maximizes the correlation among features extracted from different models but also preserves their class discrimination capability (discriminative power). While two-view methods such as CCA and DCCA enhanced individual models, multi-view approaches such as MCCA and especially DMCCA provided a more comprehensive modeling of complex inter-feature relationships across three sources. As a result, feature fusion based on DMCCA proved to be effective not only for the COWC dataset but also potentially for other remote sensing datasets. This approach shows great promise for applications requiring high accuracy and robustness, such as vehicle counting, automated urban area analysis, and satellite monitoring systems.

For future work, the DMCCA model can be further enhanced by developing its sparse or nonlinear kernel versions, as well as by integrating the learned projections into end-to-end deep learning frameworks. In this way, more flexible and efficient forms of discriminative correlation learning can be achieved.

References

[1] *Madhavan V., Nithya R., Devi G.U.* Feature fusion strategies for multimodal biometric systems: A review // Expert Systems with Applications. – 2021. – Vol. 168. – Art. 114336.

[2] *Liu Z., Blasch E.* Statistical analysis of the performance assessment results for pixel-level image fusion // 17th International Conference on Information Fusion (FUSION). – Salamanca, Spain, 2014. – P. 1-8.

[3] *Atrey P.K., Hossain M.A., El Saddik A., Kankanhalli M.S.* Multimodal fusion for multimedia analysis: a survey // Multimedia Systems. – 2010. – Vol. 16, No. 5. – P. 345-379.

[4] *Jolliffe I.* Principal component analysis // Encyclopedia of statistics in behavioral science. – 2005.
[5] *Fisher R.A.* The use of multiple measurements in taxonomic problems // Annals of eugenics. – 1936. – Vol. 7, No. 2. – P. 179-188.
[6] *Hardoon D.R., Szedmak S., Shawe-Taylor J.* Canonical correlation analysis: An overview with application to learning methods // Neural computation. – 2004. – Vol. 16, No. 12. – P. 2639-2664.
[7] *Arandjelovic O.* Discriminative extended canonical correlation analysis for pattern set matching // Machine Learning. – 2014. – Vol. 94, No. 3. – P. 353-370.
[8] *Kettenring J.R.* Canonical analysis of several sets of variables // Biometrika. – 1971. – Vol. 58, No. 3. – P. 433-451.
[9] *Gao L., Qi L., Chen E., Guan L.* Discriminative multiple canonical correlation analysis for information fusion // IEEE Transactions on Image Processing. – 2017. – Vol. 27, No. 4. – P. 1951-1965.
[10] *Mundhenk T.N., Konjevod G., Sakla W.A., Boakye K.* A Large Contextual Dataset for Classification, Detection and Counting of Cars with Deep Learning // Computer Vision – ECCV 2016 / Ed. by B. Leibe, J. Matas, N. Sebe, M. Welling. – Cham : Springer, 2016. – Vol. 9907. – (Lecture Notes in Computer Science). – P. 785-800. – DOI: https://doi.org/10.1007/978-3-319-46487-9_48.

КОРРЕЛЯЦИОННОЕ ОБЪЕДИНЕНИЕ ПРИЗНАКОВ ДЛЯ КЛАССИФИКАЦИИ ОБЪЕКТОВ НА СПУТНИКОВЫХ ИЗОБРАЖЕНИЯХ

[1*]***Фазылов Ш.Х.,*** [2]***Юсупов О.Р.,*** [1]***Хандамов Й.Х.***

*sh.fazilov@gmail.com
[1]Научно-исследовательский институт развития цифровых технологий и искусственного интеллекта, Ташкент, Узбекистан.
[2]Самаркандский государственный университет имени Шарофа Рашидова, Самарканд, Узбекистан.

Аннотация. В данном исследовании анализируются методы объединения признаков, основанные на сверточных нейронных сетях (CNN) и каноническом корреляционном анализе, с целью повышения эффективности обнаружения и классификации транспортных средств на данных дистанционного зондирования. Экспериментальные исследования были проведены на наборе данных Cars Overhead With Context (COWC). На первом этапе была выполнена тонкая настройка (fine-tuning) моделей ResNet18, DenseNet121 и VGG16 с использованием 20% набора данных, после чего из их выходных слоёв были извлечены векторные признаки. Каждая модель была оценена как индивидуально, так и в составе объединённых признаков, полученных с помощью различных статистических подходов: CCA (Canonical Correlation Analysis), DCCA (Discriminative Canonical Correlation Analysis), MCCA (Multiple Canonical Correlation Analysis) и DMCCA (Discriminative Multiple Canonical Correlation Analysis). Результаты показали, что среди отдельных CNN-моделей наивысшую точность продемонстрировала модель DenseNet121 (≈87.45%). Однако объединённые при помощи DMCCA признаки, классифицированные с использованием SVM, обеспечили лучшую общую производительность — около 91.52% точности, а также превосходящие показатели precision, recall и F1-score.

Ключевые слова: Дистанционное зондирование, COWC, Дискриминативный множественный канонический корреляционный анализ, DMCCA, Объединение признаков, SVM, Классификация.

SUN’IY YO‘LDOSH TASVIRLARIDA OBYEKTLARNI TASNIFLASHDA KORRELYATSIYAGA ASOSLANGAN BELGI BIRLASHTIRISH

[1*]***Fazilov Sh.X.,*** [2]***Yusupov O.R.,*** [1]***Xandamov Y.X.***

*sh.fazilov@gmail.com
[1]Raqamli texnologiyalar va sun’iy intellektni rivojlantirish ilmiy-tadqiqot instituti, Toshkent, O‘zbekiston.
[2]Sharof Rashidov nomidagi Samarqand davlat universiteti, Samarqand, O‘zbekiston.

Annotatsiya. Ushbu tadqiqotda masofadan zondlash tasvirlarida avtomobillarni aniqlash va tasniflash samaradorligini oshirish maqsadida konvolyutsion neyron tarmoqlar (CNN) va kanonik korrelyatsiya tahliliga asoslangan belgilarni birlashtirish usullari tahlil qilindi. Tajribaviy ishlar Cars Overhead With Context (COWC) datasetida olib borildi. Dastlab ResNet18, DenseNet121 va VGG16 modellarida 20% ma’lumot asosida fine-tuning jarayoni o‘tkazilib, ularning chiqish qatlamlaridan olingan belgilar (feature vektorlar) ajratib olindi. Har bir model alohida holda va ularning belgilarini turli statistik usullari – CCA (Canonical Correlation Analysis), DCCA (Discriminative Canonical Correlation Analysis), MCCA (Multiple Canonical Correlation Analysis) hamda DMCCA (Discriminative Multiple Canonical Correlation Analysis) yordamida birlashtirish orqali sinovdan o‘tkazildi. Natijalar shuni ko‘rsatdiki, alohida CNN modellar ichida DenseNet121 eng yuqori aniqlikka (≈ 87.45%) ega bo‘ldi, biroq DMCCA yordamida olingan birlashtirilgan belgilar SVM tasniflagichi bilan qo‘llanilganda eng yuqori aniqlik (≈ 91.52%), accuracy, precision, Recall va F1-score ko‘rsatkichlariga erishildi.

Kalit so‘zlar: Masofadan zondlash, COWC, Discriminative Multiple Canonical Correlation Analysis, DMCCA, Feature Fusion, SVM, Tasniflash.