

Bag-of-visual word (BoVW) merupakan salah satu pendekatan atau model yang banyak digunakan untuk mengklasifikasi sebuah citra. Model BoVW mengubah citra dalam bentuk vektor yang tersusun dari banyaknya *visual word* pada citra tersebut. *Visual word* merupakan pusat dari sekelompok *keypoint* yang diekstrak dari banyak citra. *Keypoint-keypoint* tersebut terletak pada tepi objek pada citra dan mampu memberikan deskripsi pada citra tersebut. Performa klasifikasi citra berbasis BoVW bergantung pada kualitas *visual word* yang dihasilkan. Rendahnya kualitas *visual word* dapat dipengaruhi oleh (i) pemberian bobot yang belum optimal dan (ii) banyaknya jumlah *keypoint* yang dihasilkan. Selain berdampak pada rendahnya akurasi, banyaknya jumlah *keypoint* juga menyebabkan lamanya proses pembentukan *visual word* pada fase *training* dalam proses klasifikasi.

Beberapa skema pembobotan fitur global telah dikembangkan untuk meningkatkan kemampuan fitur dalam proses klasifikasi. Salah satunya adalah Inverse Document Frequency (IDF) yang merupakan skema pembobotan global berbasis Document Frequency (DF). Selain banyak digunakan pada klasifikasi teks, skema pembobotan tersebut juga digunakan pada klasifikasi citra. Namun kelemahan skema pembobotan fitur yang berbasis DF adalah rendahnya kemampuan diskriminatif fitur dalam proses klasifikasi, sehingga penelitian ini mengusulkan skema pembobotan global Term Distribution (TD) yang menggunakan distribusi *intra-class* dan *inter-class* dari sebuah *visual word*. Banyaknya jumlah *keypoint* yang dihasilkan dapat diatasi dengan menggunakan metode pemilihan *keypoint*. Beberapa metode telah diusulkan yaitu Iterative Keypoint Selection (IKS1 dan IKS2). Kedua metode tersebut menggunakan inisialisasi *keypoint* yang dilakukan secara random, sehingga menyebabkan hasil klasifikasi yang berubah-ubah. Untuk mengatasi hal tersebut, penelitian ini mengusulkan metode pemilihan *keypoint* yang berbasis *distance matrix* atau Distance Matrix-based Keypoint Selection (DMKS).

Eksperimen dilakukan dengan menggunakan dataset publik, Coil-100, Caltech-101, dan Caltech-256. Hasil pengujian menunjukkan metode yang diusulkan mampu menghasilkan performa yang lebih baik dibandingkan dengan metode sebelumnya. Skema pembobotan fitur TD mampu meningkatkan akurasi klasifikasi dengan selisih hingga 43,6% dan 31,54%, lebih baik dibandingkan mIDF dan IGM. Penggunaan distribusi *visual word* pada skema pembobotan global mampu meningkatkan kemampuan diskriminatif *visual word* dalam proses klasifikasi. Untuk percobaan pemilihan *keypoint*, DMKS mampu menghasilkan akurasi yang lebih baik dari IKS1 dan IKS2 dengan selisih akurasi sebesar 0,84% dan 0,48%. Walaupun selisih akurasi yang dihasilkan DMKS tidak lebih dari 1%, namun waktu komputasinya dapat lebih cepat sekitar 20 kali waktu komputasi IKS1 dan 10 kali waktu komputasi IKS2 pada pengujian menggunakan dataset Caltech-101. Penggunaan *distance matrix* pada DMKS dapat mempercepat proses pemilihan *keypoint* dibandingkan pendekatan *iterative* pada IKS1 dan IKS2. Dari hasil percobaan tersebut, metode usulan dapat dimanfaatkan untuk meningkatkan performa klasifikasi citra, terutama klasifikasi citra berbasis model BoVW dengan objek citra yang berbeda seperti klasifikasi citra satelit, citra medis, dan citra bawah air.

Kata kunci – *Bag-of-visual words*, klasifikasi citra, pembobotan fitur, seleksi *keypoint*

The bag-of-visual words (BoVW) model has been widely used for image classification. This model represents an image into a vector that contains the number of visual words. The visual words are the center of grouped keypoints, which are extracted from the images. These keypoints are located at the edge of an object and capable to describe the image. The performance of BoVW based image classification depends on the quality of the visual words. The low quality of visual words can be influenced by: (i) weighting scheme is not optimal and (ii) the large number of generated keypoints. Not only produce a low accuracy of image classification but also lead to the high computational time of visual words generation.

Several global weighting schemes have been developed to improve the performance of classification. There is Inverse Document Frequency (IDF) as a global weighting scheme based on Document Frequency (DF). Weighting schemes have been used not only in text classification but also in image classification. Global weighting schemes based on DF are not sensitive enough to measure the discriminating power of visual words. Therefore, this study introduces a global weighting scheme called Term Distribution (TD) based on the intra-class and inter-class distributions of the visual words instead of DF. To reduce the large number of keypoints, we proposed keypoint selection which is based on the distance matrix, namely Distance Matrix-based Keypoint Selection (DMKS). The distance matrix approach addresses the randomized keypoint initialization problem of previous methods.

The experiments were carried out using public datasets, i.e. Coil-100, Caltech-101, and Caltech-256. The results show that the proposed methods can produce better performance compared to the previous methods. Our proposed weighting scheme achieved better accuracies than previous weighting schemes. The difference accuracy is up to 43.6 % and 31.54 %. The use of the visual words distribution in global weighting can increase the discriminative power of visual words in the classification. For the keypoint selection, DMKS was able to produce better accuracies than IKS1 and IKS2 with an accuracy difference of 0.84 % and 0.48 %. Although the difference in accuracies generated by DMKS is not more than 1%, the computation time can be faster, around 20 times in the IKS1 and 10 times in the IKS2 using the Caltech-101 dataset. The use of distance matrix in DMKS can speed up the process of selecting keypoints compared to the iterative approach in IKS1 and IKS2. Our results indicate that the proposed methods can be utilized to improve the performance of image classification, especially BoVW based image classification for another object, such as satellite imagery, medical imagery, and underwater imagery.

Keywords – Bag-of-visual words, image classification, weighting schemes, keypoint selection