Supplementary materials

1. Pseudocode of the proposed method

The pseudocodes of the training and the test phases are described in Fig. 1.

```
Function Training_Phase(training_data)
  (Feature space selection)
  feature_space is selected in training_data by feature space selection
  for each class-pair CP_ij
    (Feature vector selection)
    for each location l ∈ feature_space
      for each searchlight window w
        perf(w) = leave-one-trial-out cross-validation performance of CP_ij using a
                  feature vector inside w centered on l
      end
      best_window(CP_ij, l) = argmax_w perf(w)
      reliability(CP_ij, l) = perf(best_window(CP_ij, l))
    end
    for k = 1 to maximum number of sub-classifiers
      ensemble_perf(k) = leave-two-trial-out cross-validation performance of CP_ij
                         using k most reliable feature vectors
    end
    M(CP_ij) = argmax_k ensemble_perf (k)
    (Multiple sub-classifiers)
    for n = 1 to M(CP_ij)
      The sub-classifier SC(CP_ij, n) is constructed with n-th reliable feature vector
      based on reliability(CP_ij, l) and best_window(CP_ij, l).
    end
  end
Return SC, M
```

```
Function Test_Phase(test_data, SC, M)
  for each class-pair CP_ij
    (Multiple sub-classifiers)
    for n = 1 to M(CP_ij)
      P(CP_ij, i, n) = probability that test_data is class i, estimated by SC(CP_ij, n)
      P(CP_ij, j, n) = probability that test_data is class j, estimated by SC(CP_ij, n)
    end
    (Voting values for two classes)
    for each class-pair CP_ij
      V(CP_ij, i) = Σ_n P(CP_ij, i, n) / M(CP_ij)
      V(CP_ij, i) = Σ_n P(CP_ij, i, n) / M(CP_ij)
    end
  end
  for each class i
    sum_vote(i) = sum of vote for i from all class-pair related to i
  end
  Estimated_Class = argmax_i sum_vote(i)
Return Estimated_Class
```

2. Feature space selection methods

The five feature selection methods were as follows:

a) Whole brain: All of the voxels in the brain region were selected as the feature space.

b) Analysis of variance (ANOVA): ANOVA evaluates the usefulness of each voxel by measuring the consistency of intensity variation with respect to stimuli. Using the ANOVA criteria, the most informative 1000 voxels were selected to be used as the feature space.

c) Support vector machine (SVM): Initially the voxels are ranked according to the absolute values of the weights, and subsequently, the most influential voxels in the construction of the decision boundary are identified[18]. Finally, a total of 1000 voxels were selected from among the classifiers for every class.

d) Recursive feature elimination (RFE): In the RFE, the voxels that are estimated to have less information are eliminated from the feature space through iterations[18]. In each iteration, the voxels were ranked and selected using the same method as that for SVM feature selection, and 5% of the less informative voxels were eliminated at each iteration until the number of remaining voxels reached 1000.

e) Ventral temporal cortex (VTC) mask: The VTC region of the brain is known to be responsible for the classification of visually-presented objects. In this paper, we applied VTC masks provided along with the dataset (details refer to 2.2.2. fMRI dataset). For inter-subject analysis, the mask for one subject was registered into the standard space.

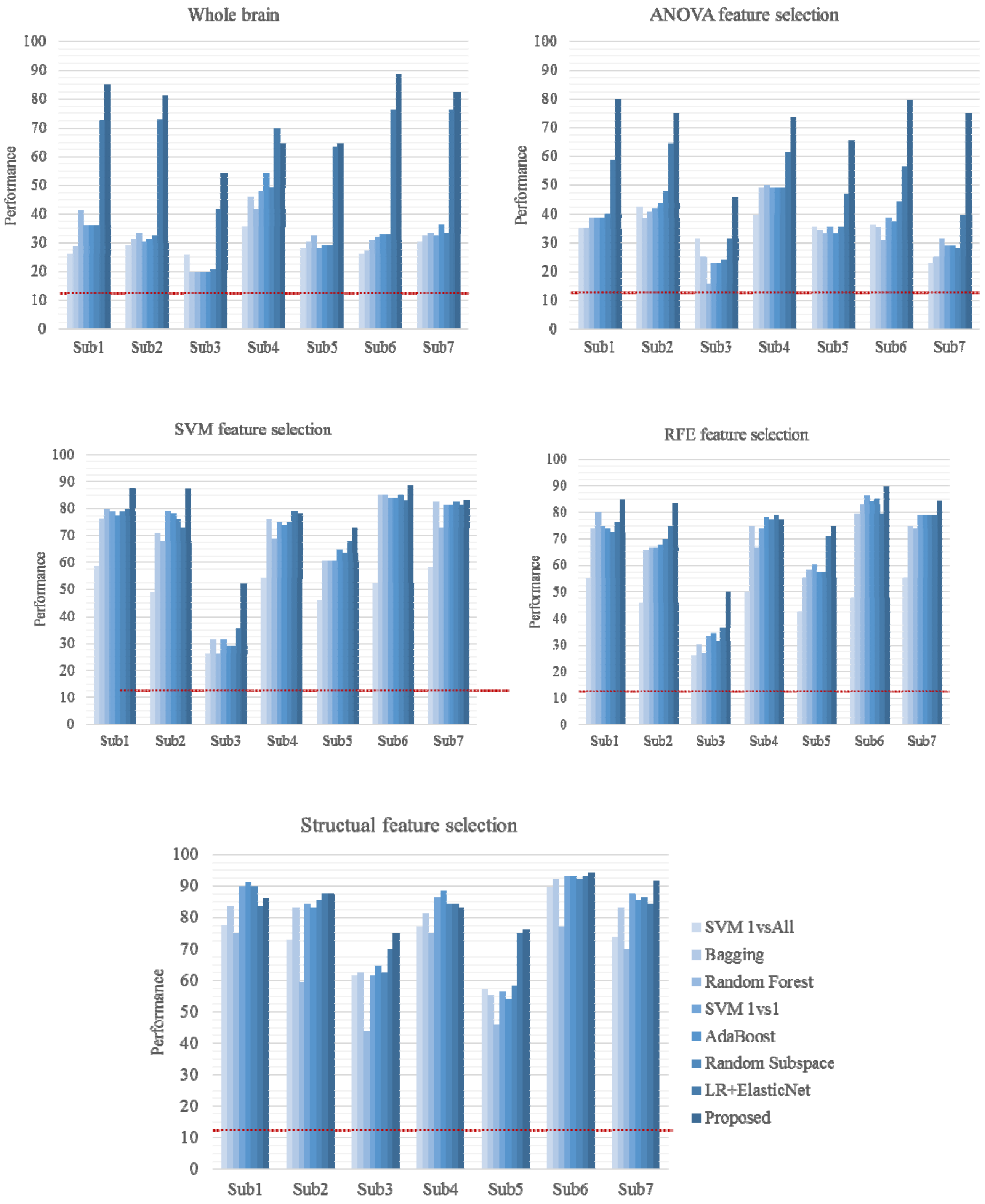

Fig. S1. The detailed performances for individual intra-subject classification. The dotted line indicates the chance level of 12.5% for eight classes.