# "Multi-objective Differential Evolution for Feature Selection in Classification" Online Supplementary Materials 

Peng Wang, Student Member, IEEE, Bing Xue, Senior Member, IEEE, Jing Liang, Senior Member, IEEE, and Mengjie Zhang, Fellow, IEEE

## I. Introduction

This is the Online Supplementary Materials of Multiobjective Differential Evolution for Feature Selection in Classification.

Tables I and II show the average test classification accuracy and the average number of selected features of the solution with the best training accuracy under the results of 30 runs. Table III presents the average $P V$ and $A N$ results of Omnioptimizer, DN-NSGAII, MO_Ring_PSO, and MOCDE. Table IV gives some examples of different feature subsets with the same objective values on the Multiple, Arrhythmia, SRBCT, Leukemia, and DLBCL datasets.

To readers' convenience, an example of the representation of individuals is given in Fig. 1. Meanwhile, Fig. 2 presents the performance of the proposed initialization method (PIM) on two multi-objective optimization methods, i.e., NSGA-II and SPEA2.

To illustrate the effect of the frequency of grouping $m$ on the algorithm performance, the average results of $I_{H}$ and $I_{I G D}$ according to the different values of $m$ are shown in Fig. S.3. Five situations under four datasets (Sonar, Musk1, SRBCT and Leukemia datasets) where $m=1,5,10,15,20$ (scale value in horizontal axes) are explored, respectively. To obtain the $I_{I G D}$ results on the test sets, the true PF is still estimated by the nondominated solutions obtained from the eight algorithms with 30 independent runs.

## II. The Representation of Individuals

In Fig. 1, suppose that a dataset includes $n$ features, $F_{1}$ to $F_{n}$, and the individuals $S_{1}$ to $S_{P}$ form a population with size $P$. Each individual in DE is a $n$-dimensional vector, and one dimension corresponds to one feature. Each element in an individual is a real number between 0 and 1 . If a value is not less than a predefined threshold $\theta$, e.g., $\theta=0.6$, its corresponding feature will be selected. Otherwise, it will not be selected. For example, the first feature $F_{1}$ is not selected in $S_{1}$, but selected in $S_{2}$.

[^0]

Fig. 1. An example of the encoding scheme for the individuals in different spaces.


Fig. 2. The $I_{H}$ and $I_{I G D}$ of NSGA-II and SPEA2 with and without the proposed PIM initialization method.

## III. Analysis on the Initialization and Clustering INTERVAL

## A. Performance of the proposed initialization method

In Fig. 2, NSGA-II or SPEA2 without PIM means NSGAII or SPEA2 uses the traditional random initialization method,


Fig. 3. The boxplots of two indicators on the test sets based on the different values of $m$.
where all features have the same probability to be selected in the initialization step. Noted that the smaller $I_{I G D}$ values mean the better algorithm performance, which is opposite for $I_{H}$. It can be observed that NSGA-II with PIM achieves larger $I_{H}$ values and smaller $I_{I G D}$ values on almost all the 14 datasets compared with NSGA-II with the traditional initialization method.

As for SPEA2, by using the proposed PIM initialization method, SPEA2 obtains a higher $I_{H}$ value than the traditional initialization method apart from the Wine dataset. The superiority is particularly obvious on the last nine datasets. A similar trend can be seen on the $I_{I G D}$ results.

The results indicate that on almost all the 14 datasets, applying the MIC-based initialization method can help EMO-based feature selection algorithms to achieve better classification performance.

## B. Analysis on the effect of the clustering interval $m$

In Fig. 3, the average $I_{H}$ and $I_{I G D}$ on the test sets have little difference when the value of $m$ of a dataset is different. The average difference does not exceed 0.1 for both $I_{H}$ and $I_{I G D}$. It stays almost unchanged, e.g., the average $I_{I G D}$ on the Musk1 dataset and the average $I_{H}$ on the Sonar dataset. But there are still some fluctuations, which are more obvious when the number of features is large. For example, for the Leukemia dataset, the largest value of average $I_{I G D}$ is 0.267 while the smallest value is 0.012 .

As mentioned above, the larger the $I_{H}$ and the lower the $I_{I G D}$, the better the algorithm. When $m=1$, the upper limit of $I_{H}$ and the lower limit of $I_{I G D}$ are not the highest and lowest compared with other values of $m$ in any of those four datasets. Also, there are more large outliers for average $I_{I G D}$ on the SRBCT and Leukemia datasets than that of $m=5,10,15$.

According to the average values of $I_{H}$ and $I_{I G D}$ in Fig. 3, the proposed MOCDE method is not very sensitive to the values of the parameter $m$. Cluster interval $m=1$ means that MOCDE will perform clustering at each generation, which increases the computational complexity of the
proposed MOCDE algorithm. When $m$ is too big, e.g., 20, the performance is slightly decreased in general. Therefore, a smaller value of $m$ is not recommended. Based on Fig. 3, $m=10$ is a good starting point.

## IV. COMPUTATIONAL COMPLEXITY ANALYSIS

In this subsection, the computation complexity of the used dynamic SCD and the hypervolume contribution indicator is analyzed. Then, the total computation complexity of MOCDE is discussed.

Assume using a population with its size of $N$ to solve a problem with $M$ objectives and $n$ decision variables. The computational complexity of traditional crowding distance is $O(M N \log N)$ [1]. The difference in computation complexity between SCD and the traditional crowding distance lies in the calculation of crowding distance in the search space. Then, the computation complexity of SCD is $O((M+n) N \log N)$. Suppose that the number of solutions that need to be removed is $t$ in Algorithm 2 of Section III.F, the computational complexity of the used dynamic SCD is $O((M+n+t) N \log N)$. Similarly, the computational complexity of the used dynamic hypervolume contribution indicator is $O\left((m+N)^{M}\right)$ ( $m$ is the number of solutions that need to be removed in Algorithm 3 of Section III.G), and that of the original hypervolume contribution indicator is $O\left(N^{M}\right)$ [2]. In this work, $M$ is 2 , and the minimum and the maximum values of $t$ and $m$ are 0 and $N$. Therefore, the computation complexity of the dynamic SCD is in the range of $\left[O(n N \log N), O\left(n N \log N^{2}\right)\right]$. Similarly, the computational complexity of the used dynamic hypervolume contribution indicator in the proposed MOCDE algorithm is between $O\left(N^{2}\right)$ and $O\left(4 N^{2}\right)$.

The complexity of the proposed MOCDE algorithm mainly includes three aspects: the main iteration loop of DE, clustering, and the calculations of SCD and $C_{H}$. The complexity of the main loop and clustering are $O(N)$ and $O\left(N^{2}\right)$. Therefore, the total computational complexity of MOCDE is between $O\left(N^{2}\right)$ and $O\left(N^{3}+n N^{2} \log N\right)$.

TABLE I: The average classification accuracy results on the test sets

| Dataset | NSGA-II | SPEA2 | MOEA/D | SparseEA | Omni-optimizer | DN-NSGAII | MO_Ring_PSO | MOCDE |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Wine | $95.83 \%$ | $96.42 \%$ | $96.09 \%$ | $\mathbf{9 6 . 7 9 \%}$ | $94.99 \%$ | $94.94 \%$ | $93.10 \%$ | $95.05 \%$ |
| Zoo | $89.19 \%$ | $86.43 \%$ | $88.72 \%$ | $84.79 \%$ | $86.77 \%$ | $86.37 \%$ | $82.69 \%$ | $\mathbf{9 0 . 0 3 \%}$ |
| SPECT | $65.72 \%$ | $63.33 \%$ | $65.76 \%$ | $61.89 \%$ | $66.66 \%$ | $64.76 \%$ | $63.83 \%$ | $\mathbf{6 9 . 2 4 \%}$ |
| WBCD | $88.22 \%$ | $88.20 \%$ | $88.49 \%$ | $87.92 \%$ | $88.49 \%$ | $88.66 \%$ | $87.91 \%$ | $\mathbf{9 0 . 9 4 \%}$ |
| Ionosphere | $85.94 \%$ | $86.53 \%$ | $86.01 \%$ | $86.06 \%$ | $86.04 \%$ | $85.63 \%$ | $85.73 \%$ | $\mathbf{8 6 . 9 6 \%}$ |
| Sonar | $67.66 \%$ | $65.00 \%$ | $67.74 \%$ | $67.14 \%$ | $66.83 \%$ | $66.63 \%$ | $68.65 \%$ | $\mathbf{7 6 . 9 8 \%}$ |
| Movement | $61.96 \%$ | $63.42 \%$ | $62.33 \%$ | $61.85 \%$ | $61.22 \%$ | $60.30 \%$ | $61.51 \%$ | $\mathbf{7 0 . 4 5 \%}$ |
| Hillvally | $53.14 \%$ | $53.25 \%$ | $53.25 \%$ | $54.34 \%$ | $51.39 \%$ | $51.87 \%$ | $51.88 \%$ | $\mathbf{5 5 . 7 2 \%}$ |
| Musk1 | $96.74 \%$ | $96.92 \%$ | $97.21 \%$ | $96.57 \%$ | $95.16 \%$ | $95.45 \%$ | $95.79 \%$ | $\mathbf{9 7 . 4 8 \%}$ |
| Multiple | $92.88 \%$ | $93.46 \%$ | $92.88 \%$ | $93.69 \%$ | $93.96 \%$ | $93.70 \%$ | $94.22 \%$ | $\mathbf{9 5 . 0 9 \%}$ |
| Arrhythmia | $\mathbf{5 7 . 5 5 \%}$ | $56.92 \%$ | $57.00 \%$ | $50.35 \%$ | $55.14 \%$ | $55.75 \%$ | $55.89 \%$ | $56.01 \%$ |
| SRBCT | $75.59 \%$ | $78.73 \%$ | $77.35 \%$ | $72.55 \%$ | $76.77 \%$ | $77.35 \%$ | $79.31 \%$ | $\mathbf{8 5 . 8 3 \%}$ |
| Leukemia | $82.63 \%$ | $81.75 \%$ | $83.57 \%$ | $85.86 \%$ | $87.37 \%$ | $86.93 \%$ | $86.93 \%$ | $\mathbf{8 9 . 5 6 \%}$ |
| DLBCL | $87.41 \%$ | $87.65 \%$ | $86.46 \%$ | $75.69 \%$ | $81.10 \%$ | $81.74 \%$ | $81.83 \%$ | $\mathbf{9 0 . 0 3 \%}$ |

TABLE II: The average size results on the test sets

| Dataset | NSGA-II | SPEA2 | MOEA/D | SparseEA | Omni-optimizer | DN-NSGAII | MO_Ring_PSO | MOCDE |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Wine | $\mathbf{3 . 5 3}$ | 4.00 | 3.98 | 4.30 | 4.97 | 4.93 | 4.37 | 4.03 |
| Zoo | 6.33 | 6.27 | 6.32 | $\mathbf{5 . 7 3}$ | 6.97 | 7.23 | 6.87 | 7.43 |
| SPECT | 7.07 | 7.07 | 6.88 | $\mathbf{4 . 5 7}$ | 6.87 | 6.93 | 7.10 | 6.83 |
| WBCD | 4.83 | 4.53 | $\mathbf{4 . 2 6}$ | 7.13 | 10.23 | 10.43 | 8.37 | 8.17 |
| Ionosphere | 5.97 | $\mathbf{5 . 0 7}$ | 5.12 | 6.43 | 8.53 | 9.93 | 6.73 | 7.13 |
| Sonar | 10.50 | $\mathbf{1 0 . 2 0}$ | 11.00 | 14.87 | 20.00 | 19.80 | 18.23 | 17.33 |
| Movement | $\mathbf{1 4 . 4 7}$ | 14.87 | 14.68 | 20.10 | 29.57 | 32.60 | 29.40 | 22.67 |
| Hillvally | 13.77 | 12.23 | 12.11 | $\mathbf{1 1 . 4 0}$ | 34.00 | 33.50 | 30.00 | 19.57 |
| Musk1 | 17.27 | 15.83 | 16.57 | $\mathbf{1 3 . 9 3}$ | 57.03 | 55.83 | 51.77 | 38.10 |
| Multiple | $\mathbf{4 7 . 0 7}$ | 58.73 | 66.35 | 76.57 | 89.40 | 86.93 | 96.43 | 101.60 |
| Arrhythmia | 27.73 | 28.00 | 26.52 | $\mathbf{1 1 . 1 7}$ | 95.87 | 97.47 | 97.33 | 63.47 |
| SRBCT | 100.20 | 117.40 | 105.24 | $\mathbf{1 1 . 8 0}$ | 826.63 | 831.53 | 832.37 | 51.23 |
| Leukemia | 565.57 | 594.70 | 535.87 | $\mathbf{4 . 0 0}$ | 1912.57 | 1918.60 | 1871.30 | 35.93 |
| DLBCL | 1031.20 | 1107.53 | 857.25 | $\mathbf{8 . 9 0}$ | 2701.23 | 2707.30 | 2648.17 | 125.40 |

TABLE III: The $P V$ and $A N$ of the four algorithms (Omni-optimizer, DN-NSGAII, MO_Ring_PSO, and MOCDE)

| Dataset | Omni-optimizer |  | DN-NSGAII |  | MO_Ring_PSO |  | MOCDE |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $P V$ | $A N$ | $P V$ | $A N$ | $P V$ | $A N$ | $P V$ | $A N$ |
| Wine | $0 \%$ | 0 | $0 \%$ | 0 | $6.7 \%$ | 0.1 | $\mathbf{9 6 . 7 \%}$ | $\mathbf{2 . 3}$ |
| Zoo | $13.3 \%$ | 0.4 | $20 \%$ | 0.4 | $80 \%$ | 4.0 | $\mathbf{1 0 0 \%}$ | $\mathbf{6 . 1}$ |
| SPECT | $16.7 \%$ | 0.4 | $6.7 \%$ | 0.1 | $86.7 \%$ | 3.3 | $\mathbf{1 0 0 \%}$ | $\mathbf{1 4 . 7}$ |
| WBCD | $6.7 \%$ | 0.1 | $20 \%$ | 0.5 | $26.7 \%$ | 0.6 | $\mathbf{1 0 0 \%}$ | $\mathbf{6 . 0}$ |
| Ionosphere | $3.3 \%$ | 0.1 | $10 \%$ | 0.2 | $36.7 \%$ | 1.0 | $\mathbf{9 3 . 3 \%}$ | $\mathbf{7 . 5}$ |
| Sonar | $3.3 \%$ | 0.1 | $6.7 \%$ | 0.1 | $20 \%$ | 0.4 | $\mathbf{9 6 . 7 \%}$ | $\mathbf{1 0 . 1}$ |
| Movement | $13.3 \%$ | 0.3 | $16.7 \%$ | 0.3 | $3.3 \%$ | 0.1 | $\mathbf{9 3 . 3 \%}$ | $\mathbf{6 . 3}$ |
| Hillvally | $3.3 \%$ | $\mathbf{0 . 1}$ | $\mathbf{6 . 7 \%}$ | $\mathbf{0 . 1}$ | $3.3 \%$ | $\mathbf{0 . 1}$ | $0 \%$ | 0 |
| Musk1 | $80 \%$ | 3.3 | $83.3 \%$ | 3.4 | $10 \%$ | 0.2 | $\mathbf{9 3 . 3 \%}$ | $\mathbf{1 6 . 4}$ |
| Multiple | $\mathbf{1 0 0 \%}$ | 10.8 | $\mathbf{1 0 0 \%}$ | 12.4 | $0 \%$ | 0 | $\mathbf{1 0 0 \%}$ | $\mathbf{2 7 . 9}$ |
| Arrhythmia | $90 \%$ | 6.3 | $76.7 \%$ | $\mathbf{6 . 4}$ | $0 \%$ | 0 | $\mathbf{1 0 0 \%}$ | 2.9 |
| SRBCT | $90 \%$ | 25.3 | $93.3 \%$ | $\mathbf{3 6 . 5}$ | $23.3 \%$ | 0.5 | $\mathbf{1 0 0 \%}$ | 8.7 |
| Leukemia | $93.3 \%$ | 38.6 | $\mathbf{1 0 0 \%}$ | $\mathbf{7 8 . 4}$ | $13.3 \%$ | 0.3 | $\mathbf{1 0 0 \%}$ | 10.3 |
| DLBCL | $\mathbf{1 0 0 \%}$ | 80.6 | $\mathbf{1 0 0 \%}$ | $\mathbf{9 1 . 6}$ | $3.3 \%$ | 0.1 | $93.3 \%$ | 15.2 |

## REFERENCES

[1] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," IEEE Trans. Evol. Comput., vol. 6, no. 2, pp. 182-197, 2002.
[2] A. P. Guerreiro, C. M. Fonseca, and L. Paquete, "The hypervolume indicator: Problems and algorithms," arXiv preprint arXiv:2005.00515, 2020.

TABLE IV: The obtained different feature subsets with the same training accuracy on the five datasets

| Dataset | Feature subsets | Accuracy |
| :---: | :---: | :---: |
| Multiple | $\left\{F_{1}, F_{3}, F_{15}, F_{19}, F_{28}, F_{30}, F_{32}, F_{34}, F_{45}, F_{48}, F_{51}, F_{55}, F_{56}, F_{64}, F_{68}, F_{73}, F_{75}, F_{80}, F_{87}\right.$, $F_{103}, F_{105}, F_{109}, F_{110}, F_{112}, F_{115}, F_{122}, F_{139}, F_{141}, F_{150}, F_{155}, F_{158}, F_{163}, F_{168}, F_{180}, F_{185}, F_{188}, F_{190}, F_{192}, F_{193}$, $\left.F_{197}, F_{202}, F_{206}, F_{213}, F_{218}, F_{221}, F_{224}, F_{229}, F_{233}, F_{238}, F_{240}\right\}$, $\left\{F_{12}, F_{19}, F_{25}, F_{29}, F_{30}, F_{39}, F_{48}, F_{55}, F_{56}, F_{58}, F_{65}, F_{66}, F_{68}, F_{70}, F_{72}, F_{79}, F_{83}, F_{86}, F_{88}, F_{91}\right.$, $F_{94}, F_{97}, F_{101}, F_{103}, F_{106}, F_{110}, F_{112}, F_{116}, F_{125}, F_{126}, F_{129}, F_{134}, F_{138}, F_{139}, F_{146}, F_{152}, F_{159}, F_{161}, F_{163}, F_{169}$, $\left.F_{170}, F_{172}, F_{183}, F_{193}, F_{202}, F_{206}, F_{218}, F_{221}, F_{226}, F_{231}, F_{233}\right\}$ | 97.0\% |
|  | $\left\{F_{2}, F_{3}, F_{4}, F_{8}, F_{12}, F_{15}, F_{16}, F_{18}, F_{29}, F_{30}, F_{38}, F_{40}, F_{41}, F_{44}, F_{48}, F_{51}, F_{56}, F_{59}, F_{61}, F_{63}, F_{65}\right.$, <br> $F_{67}, F_{71}, F_{72}, F_{73}, F_{75}, F_{80}, F_{81}, F_{82}, F_{94}, F_{96}, F_{97}, F_{100}, F_{102}, F_{103}, F_{108}, F_{109}, F_{110}, F_{113}, F_{115}, F_{119}, F_{122}$, <br> $F_{123}, F_{125}, F_{127}, F_{128}, F_{129}, F_{132}, F_{133}, F_{137}, F_{140}, F_{141}, F_{142}, F_{146}, F_{147}, F_{151}, F_{152}, F_{154}, F_{155}, F_{157}, F_{166}, F_{168}$, <br> $\left.F_{169}, F_{171}, F_{172}, F_{175}, F_{177}, F_{183}, F_{195}, F_{197}, F_{200}, F_{204}, F_{206}, F_{215}, F_{219}, F_{221}, F_{225}, F_{230}, F_{233}, F_{238}\right\}$, <br> $\left\{F_{3}, F_{7}, F_{12}, F_{14}, F_{15}, F_{18}, F_{19}, F_{20}, F_{21}, F_{26}, F_{30}, F_{34}, F_{35}, F_{39}, F_{42}, F_{44}, F_{49}, F_{51}, F_{53}, F_{56}\right.$, <br> $F_{57}, F_{66}, F_{70}, F_{75}, F_{76}, F_{77}, F_{79}, F_{80}, F_{82}, F_{86}, F_{88}, F_{91}, F_{96}, F_{97}, F_{100}, F_{103}, F_{105}, F_{108}, F_{111}, F_{112}$, <br> $F_{117}, F_{119}, F_{120}, F_{124}, F_{125}, F_{127}, F_{129}, F_{131}, F_{140}, F_{141}, F_{148}, F_{151}, F_{153}, F_{155}, F_{156}, F_{157}, F_{158}, F_{164}, F_{168}, F_{171}$, <br> $\left.F_{172}, F_{176}, F_{183}, F_{194}, F_{195}, F_{202}, F_{203}, F_{208}, F_{210}, F_{212}, F_{214}, F_{218}, F_{219}, F_{223}, F_{225}, F_{226}, F_{227}, F_{230}, F_{231}, F_{239}\right\}$ | 98.7\% |
| Arrhythmia | $\left\{F_{50}, F_{95}, F_{107}, F_{112}, F_{131}, F_{160}, F_{171}, F_{231}, F_{241}, F_{254}\right\},\left\{F_{25}, F_{50}, F_{107}, F_{112}, F_{115}, F_{123}, F_{131}, F_{160}, F_{231}, F_{241}\right\}$ | 56.5\% |
|  | $\begin{gathered} \left\{F_{4}, F_{12}, F_{15}, F_{25}, F_{32}, F_{33}, F_{34}, F_{39}, F_{41}, F_{46}, F_{49}, F_{56}, F_{59}, F_{80}, F_{81}, F_{91}, F_{96}, F_{97}, F_{105}, F_{109},\right. \\ F_{115}, F_{120}, F_{123}, F_{128}, F_{137}, F_{151}, F_{168}, F_{171}, F_{174}, F_{178}, F_{181}, F_{207}, F_{211}, F_{212}, F_{215}, F_{227}, F_{228}, F_{238}, F_{245}, F_{248}, \\ \left.F_{259}, F_{260}, F_{275}\right\} \\ \left\{F_{15}, F_{32}, F_{36}, F_{37}, F_{39}, F_{49}, F_{54}, F_{59}, F_{62}, F_{74}, F_{75}, F_{84}, F_{87}, F_{89}, F_{91}, F_{94}, F_{103}, F_{107}, F_{112}, F_{114},\right. \\ F_{122}, F_{125}, F_{129}, F_{137}, F_{138}, F_{143}, F_{149}, F_{155}, F_{156}, F_{157}, F_{165}, F_{167}, F_{175}, F_{178}, F_{179}, F_{192}, F_{198}, F_{205}, F_{235}, F_{237}, \\ \left.F_{241}, F_{250}, F_{256}, F_{270}\right\} \end{gathered}$ | 63.5\% |
| SRBCT | $\left\{F_{159}, F_{169}, F_{236}, F_{246}, F_{256}, F_{337}, F_{391}, F_{392}, F_{509}, F_{512}, F_{577}, F_{622}, F_{685}, F_{718}, F_{754}, F_{816}, F_{870}, F_{926}, F_{975}, F_{1008}\right.$ $\left.F_{1044}, F_{1071}, F_{1232}, F_{1256}, F_{1305}, F_{1370}, F_{1455}, F_{1475}, F_{1575}, F_{1611}, F_{1635}, F_{1731}, F_{1747}, F_{1949}, F_{2009}, F_{2078}, F_{2079}, F_{2121}\right\}$ $\left\{F_{5}, F_{59}, F_{151}, F_{169}, F_{237}, F_{244}, F_{246}, F_{375}, F_{391}, F_{411}, F_{509}, F_{521}, F_{572}, F_{621}, F_{754}, F_{778}, F_{815}, F_{816}, F_{964}, F_{988}\right.$, $\left.F_{1001}, F_{1358}, F_{1381}, F_{1443}, F_{1475}, F_{1559}, F_{1578}, F_{1603}, F_{1611}, F_{1809}, F_{1953}, F_{2008}, F_{2014}, F_{2084}, F_{2140}, F_{2175}, F_{2210}, F_{2252}\right\}$ | 96.1\% |
|  | $\left\{F_{93}, F_{157}, F_{182}, F_{238}, F_{257}, F_{261}, F_{276}, F_{335}, F_{349}, F_{361}, F_{482}, F_{509}, F_{545}, F_{566}, F_{738}, F_{744}, F_{745}, F_{782}, F_{791}, F_{829}\right.$, $F_{830}, F_{869}, F_{937}, F_{951}, F_{975}, F_{978}, F_{1057}, F_{1074}, F_{1269}, F_{1274}, F_{1370}, F_{1410}, F_{1431}, F_{1503}, F_{1596}, F_{1601}, F_{1694}, F_{1778}, F_{1835}, F_{1856}$, $\left.F_{1935}, F_{1955}, F_{1957}, F_{2053}, F_{2069}, F_{2084}, F_{2292},\right\}$ $\left\{F_{93}, F_{136}, F_{149}, F_{238}, F_{248}, F_{261}, F_{361}, F_{363}, F_{499}, F_{509}, F_{545}, F_{566}, F_{623}, F_{705}, F_{744}, F_{745}, F_{830}, F_{869}, F_{937}, F_{951}\right.$, $F_{975}, F_{978}, F_{1057}, F_{1067}, F_{1254}, F_{1410}, F_{1431}, F_{1441}, F_{1524}, F_{1551}, F_{1595}, F_{1601}, F_{1650}, F_{1670}, F_{1675}, F_{1679}, F_{1850}, F_{1856}, F_{1939}, F_{2047}$, $\left.F_{2055}, F_{2084}, F_{2170}, F_{2208}, F_{2229}, F_{2230}, F_{2281}\right\}$ | 100\% |
| Leukemia | $\left\{F_{17}, F_{84}, F_{102}, F_{546}, F_{558}, F_{1011}, F_{1293}, F_{1333}, F_{1450}, F_{1726}, F_{1753}, F_{1819}, F_{2014}, F_{2728}, F_{2906}, F_{2913}, F_{2915}, F_{3031}, F_{3069}\right.$, $F_{3071}, F_{3184}, F_{3200}, F_{3582}, F_{3718}, F_{3726}, F_{3760}, F_{3872}, F_{4026}, F_{4033}, F_{4185}, F_{4226}, F_{4231}, F_{4306}, F_{4484}, F_{4593}, F_{4626}, F_{4645}, F_{4834}$, $\left.F_{4880}, F_{4938}, F_{5087}\right\}$, $\left\{F_{678}, F_{740}, F_{985}, F_{1232}, F_{1280}, F_{1304}, F_{1349}, F_{1375}, F_{1413}, F_{1437}, F_{1478}, F_{1628}, F_{1682}, F_{1824}, F_{1904}, F_{1960}, F_{2452}, F_{2756}, F_{2816}, F_{3069}\right.$, $F_{3195}, F_{3200}, F_{3395}, F_{3465}, F_{3481}, F_{3550}, F_{3638}, F_{3789}, F_{3800}, F_{3818}, F_{3821}, F_{3859}, F_{3967}, F_{4026}, F_{4068}, F_{4119}, F_{4306}, F_{4664}$, $\left.F_{4715}, F_{4779}, F_{4947}\right\}$, $\left\{F_{160}, F_{219}, F_{266}, F_{461}, F_{530}, F_{852}, F_{1122}, F_{1184}, F_{1293}, F_{1468}, F_{1646}, F_{1732}, F_{1922}, F_{1927}, F_{1934}, F_{1960}, F_{1980}, F_{2080}, F_{2138}, F_{2158}\right.$, $F_{2502}, F_{2722}, F_{2735}, F_{2930}, F_{3045}, F_{3365}, F_{3378}, F_{3517}, F_{3728}, F_{3886}, F_{4290}, F_{4305}, F_{4497}, F_{4513}, F_{4669}, F_{4696}, F_{4713}, F_{4715}$, $\left.F_{5021}, F_{5037}, F_{5117}\right\}$ | 100\% |
| DLBCL | $\left\{F_{321}, F_{357}, F_{480}, F_{715}, F_{755}, F_{777}, F_{1247}, F_{1275}, F_{1313}, F_{1372}, F_{1705}, F_{1917}, F_{2071}, F_{2161}, F_{2452}, F_{2458}, F_{3102}, F_{3179}, F_{3187}, F_{3293}\right.$, $F_{3758}, F_{3772}, F_{3943}, F_{3949}, F_{4112}, F_{4252}, F_{4292}, F_{4298}, F_{4302}, F_{4379}, F_{4539}, F_{4621}, F_{5081}, F_{5389}, F_{5418}, F_{5437}, F_{5452}, F_{5668}, F_{5838}, F_{5936}$, $\left.F_{6043}, F_{6244}, F_{6261}, F_{6643}, F_{6835}, F_{6868}, F_{6910}, F_{6913}, F_{6948}, F_{7023}\right\}$ <br> $\left\{F_{25}, F_{41}, F_{476}, F_{544}, F_{755}, F_{760}, F_{871}, F_{957}, F_{1013}, F_{1134}, F_{1372}, F_{1393}, F_{1505}, F_{1562}, F_{1914}, F_{2245}, F_{2280}, F_{2351}, F_{2460}, F_{2575}\right.$, <br> $F_{2591}, F_{2778}, F_{2915}, F_{3794}, F_{3813}, F_{3837}, F_{3852}, F_{3870}, F_{4018}, F_{4214}, F_{4314}, F_{4367}, F_{4460}, F_{4574}, F_{4593}, F_{4621}, F_{4907}, F_{4949}, F_{4984}, F_{5011}$, <br> $\left.F_{5064}, F_{5105}, F_{5308}, F_{5321}, F_{5547}, F_{6074}, F_{6135}, F_{6191}, F_{6995}, F_{7037}\right\}$ | 96.1\% |
|  | $\left\{F_{5}, F_{82}, F_{476}, F_{869}, F_{1134}, F_{1405}, F_{1562}, F_{1611}, F_{1670}, F_{1676}, F_{1990}, F_{2043}, F_{2081}, F_{2193}, F_{2213}, F_{2419}, F_{2460}, F_{2497}, F_{2542}, F_{2595}\right.$, $F_{3008}, F_{3102}, F_{3604}, F_{3704}, F_{3716}, F_{3897}, F_{3949}, F_{3987}, F_{4314}, F_{4344}, F_{4504}, F_{4557}, F_{4564}, F_{4621}, F_{4838}, F_{5002}, F_{5169}, F_{5342}, F_{5389}, F_{5577}$, $\left.F_{5781}, F_{5904}, F_{6259}, F_{6277}, F_{6510}, F_{6547}, F_{6668}, F_{6941}, F_{7003}, F_{7017}, F_{7023}, F_{7046}\right\}$ $\left\{F_{15}, F_{321}, F_{740}, F_{777}, F_{957}, F_{1016}, F_{1059}, F_{1218}, F_{1344}, F_{1621}, F_{1768}, F_{1847}, F_{2158}, F_{2213}, F_{2419}^{\prime}, F_{2441}, F_{2460}, F_{2470}, F_{2513}, F_{2753}\right.$, $F_{2863}, F_{3198}, F_{3363}, F_{3656}, F_{3779}, F_{3949}, F_{3990}, F_{4273}, F_{4460}, F_{4504}, F_{4737}, F_{4777}, F_{5002}, F_{5112}, F_{5236}, F_{5418}, F_{5644}, F_{5777}, F_{5781}, F_{5867}$, $\left.F_{5888}, F_{6068}, F_{6074}, F_{6091}, F_{6259}, F_{6469}, F_{6556}, F_{6648}, F_{6910}, F_{6940}, F_{6990}, F_{7023}\right\}$ | 98.1\% |
|  | $\left\{F_{61}, F_{621}, F_{760}, F_{1013}, F_{1016}, F_{1134}, F_{1250}, F_{1405}, F_{1407}, F_{1421}, F_{1440}, F_{1454}, F_{1492}, F_{1496}, F_{1504}, F_{1561}, F_{1626}, F_{1768}, F_{1891}, F_{2094}\right.$, <br> $F_{2100}, F_{2277}, F_{2370}, F_{2460}, F_{2871}, F_{3020}, F_{3039}, F_{3102}, F_{3136}, F_{3268}, F_{3781}, F_{3927}, F_{3955}, F_{4179}, F_{4230}, F_{4280}, F_{4379}, F_{4504}, F_{4667}, F_{5055}$, <br> $\left.F_{5157}, F_{5169}, F_{5236}, F_{5496}, F_{5576}, F_{5615}, F_{5703}, F_{6256}, F_{6259}, F_{6272}, F_{6358}, F_{6736}, F_{6941}, F_{7023}\right\}$ <br> $\left\{F_{7}, F_{79}, F_{82}, F_{286}, F_{476}, F_{799}, F_{800}, F_{974}, F_{1562}, F_{1704}, F_{2006}, F_{2081}, F_{2158}, F_{2245}, F_{2459}, F_{2460}, F_{2795}, F_{2943}, F_{3079}, F_{3267}\right.$, <br> $F_{3388}, F_{3511}, F_{3596}, F_{3617}, F_{3704}, F_{3717}, F_{3781}, F_{3990}, F_{4160}, F_{4165}, F_{4219}, F_{4295}, F_{4352}, F_{4581}, F_{4621}, F_{4835}, F_{4891}, F_{4938}, F_{4949}, F_{5055}$, <br> $\left.F_{5092}, F_{5320}, F_{5361}, F_{5736}, F_{5897}, F_{5975}, F_{6280}, F_{6732}, F_{6736}, F_{6771}, F_{6802}, F_{6884}, F_{6892}, F_{7023}\right\}$ <br> $\left\{F_{47}, F_{79}, F_{357}, F_{504}, F_{649}, F_{727}, F_{735}, F_{755}, F_{1013}, F_{1381}, F_{1407}, F_{1568}, F_{1612}, F_{1768}, F_{1823}, F_{1825}, F_{2163}, F_{2164}, F_{2460}, F_{2466}\right.$, <br> $F_{2591}, F_{2726}, F_{2833}, F_{2875}, F_{2956}, F_{2974}, F_{3102}, F_{3155}, F_{3394}, F_{3596}, F_{3897}, F_{4186}, F_{4220}, F_{4228}, F_{4298}, F_{4302}, F_{4460}, F_{4495}, F_{4982}, F_{5170}$, <br> $\left.F_{5336}, F_{5686}, F_{5747}, F_{5762}, F_{5907}, F_{6051}, F_{6091}, F_{6132}, F_{6457}, F_{6736}, F_{6913}, F_{6948}, F_{7023}, F_{7040}\right\}$ | 100\% |


[^0]:    Manuscript received March 26, 2021; revised September 19, 2021; accepted November 2, 2021. (Corresponding author:Jing Liang)
    M. Zhang, B. Xue and P. Wang are with the Evolutionary Computation Research Group, Victoria University of Wellington, Wellington 6140, New Zealand.
    J. Liang is with the School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China.

