Bernd Hamann
Stratovan

11/16/2022

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:... We can also consider the following scenario: Outcome scores are defined for the possible outcomes/outcome pairs as follows: The score s of an outcome type t or type pair (t_1, t_2) is defined as a value in the interval $[-1, 1]$:

$$s = \begin{cases} -1 & \text{if } t = FN \\ -1 & \text{if } (t_1, t_2) = (FP, FP_{cl}) \\ 3/4 & \text{if } (t_1, t_2) = (TP, FP_{cl}) \\ 1 & \text{if } t = TN \\ 1 & \text{if } (t_1, t_2) = (TP, TP_{cl}) \end{cases}, \quad cl \in \{1, \ldots, C\}.$$

For example, if $(T_1, T_2, T_3) = (1,1,1)$ and $(R_1, R_2, R_3) = (0,0,0)$, then the overall 3-material classification outcome score will be $S = s_1 + s_2 + s_3 = -1-1-1 = -3$. ($S$ denotes the overall score, while s_i is the score for a single material classification outcome.) If $(R_1, R_2, R_3) = (1,1,1)$, then the overall score will be $S = s_1 + s_2 + s_3 = 1+1+1 = 3$. One could "normalize" this overall outcome score to the interval $[-1, 1]$ by dividing S by 3, i.e., the length of the string.

• Notes. • Distribution of classes $0, 1, .., C$ in data stream and score S: One must assume that class-0 materials are the vast majority in a data stream and that "a good" classification system nearly always identifies class-0 materials as class-0 materials. Thus, a score $S = s_1 + \ldots + s_{1000}$ will be very close to 1000, since a small number of FNs has hardly an impact on S.

! BUT THREAT DETECTION CAPABILITY IS MOST IMPORTANT. ...

OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

• Laplacian eigenfunctions and neural networks:... • Set of classes and their distribution/relative occurrences changing over time: Over time, it can be expected that some classes can be deleted from the set of threat material classes to be recognized — as they are no longer relevant. Certain "new" classes will be added to the set of threat classes to be recognized. Ideally, the necessary updates and revisions of a great classification system should be transparent to the "operator" of such a "self-adapting system." Regardless of all these considerations, the assessment of system classification performance should be INVARIANT to changing distributions/relative occurrences of classes. Furthermore, recognition performance concerning a specific class must be independent of the occurrence fraction that this class has in the data stream. In other words, it is crucial to characterize classification performance on a per-class basis. In addition, the feature characteristics of a material class could change over time. The impact of such a "material evolution" over time would result in a temporally decreasing detection rate for this material class. The deteriorating detection rate should TRIGGER A SELF-ADAPTING SYSTEM TO PERFORM "RE-OPTIMIZATION."

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

- Laplacian eigenfunctions and neural networks:... • A simple per-class performance characterization of streamed data classification: We consider a data stream with an associated class index stream consisting of index values **0, 1 and 2**. A class-0 material is a non-threat material; class-0 materials are therefore assumed to be occurring much more frequently than threat material classes 1 and 2. A material belonging to class cl can lead to only one of two classification outcomes: The material is correctly classified (T = true) or not correctly classified (F = false). We consider an example of a stream of class indices:

→ Correct, given index stream (ground truth)

0 0 0 0 1 0 0 2 0 1 0 0 0 1 1 0 0 2 2 2 0 0 0 0 2 0 2 0 0 2 0 2

→ Classification result of material stream processing

0 0 0 1 0 0 2 0 0 1 0 2 0 1 1 1 0 0 2 2 0 0 0 1 2 0 0 0 0 1 0 0

→ False material classifications (F)

T T T F F T F F T T T F T T T F T F T T T T T F T T F T T F T F

(True classifications (T) are shown as well.)

class	T	F	T/F	T/(T+F)
0	15	5	15/5	15/20
1	3	1	3/1	3/4
2	3	5	3/5	3/8
all	21	11	21/11	21/32

True and false classifications

The table (left) summarizes the results: i) When class-0 or class-1 materials are given, they are correctly classified in 75% of all instances.

...

OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

• Laplacian eigenfunctions and neural networks:... ii) When a class-2 material is given, it is correctly classified in only 37.5% of all occurrences.

iii) The combined, total correct recognition rate of threat materials (classes 1 and 2) is (3+3)/(4+8) = 50%.

→ A "more specific and detailed" characterization could be done by recording the answers to these questions:

- When classifying a material as a class-0, as a class-1 or class-2 material, is this classification correct? If the answer is "YES," we can call such a classification result a T0 (= true positive of class 0), a T1 (= true positive of class 1) or a T2 (= true positive of class 2), respectively. If the answer is "NO," we can call such a classification result an F0 (= false classification of a material as a class-0 material), an F1 (= false classification of a material as a class-1 material) or an F2 (= false classification of a material as a class-2 material, respectively. The table (left) summarizes the results for the 32-material data stream example from the previous page. The first row means: Materials of class 0 are correctly classified 15 times; they are incorrectly classified 3 times as class-1 and 2 times as class-2 materials. ...

ground truth	classification as 0	1	2
0	T0 15	F1 3	F2 2
1	F0 1	T1 3	F2 0
2	F0 4	F1 1	T2 3

Occurrences of T0, T1, T2 and F0, F1 and F2 cases in example.

■ OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

• Laplacian eigenfunctions and neural networks:... The table provided on the previous page (p. 4) is "more informative" than the table on p. 3: It provides more specific statistical information about mis-classification; when materials are mis-classified, statistical information is captured how the mis-classifications are distributed over the various classes representing mis-classifications. The rows of the table (p. 4) show that

15 of 20 materials of class 0 were correctly classified;
3 of 4 materials of class 1 were correctly classified; and
3 of 8 materials of class 2 were correctly classified.

The columns capture the information about the distribution of materials of various classes mis-classified as materials of a class cl. Again, the information/data provided in this table (p. 4) makes it possible to compute ratios/percentages/relative values that characterize system classification performance in a per-class manner — independently of the total numbers of material segments of a specific class in a data stream.

• In summary, the table (p. 4) shows which classes have a "high" PD rate (classes 0 and 1, 75%) and which have a "low" PD rate (class 2, 37.5%). More importantly, one sees that, for example, class-2 materials are often mis-classified as class-0 materials (in 50% of all cases. Thus, one must use more discriminating features for such classes.

...