Accuracy Evaluation of MEKOLAS®

1. Method for Evaluating the Accuracy of MEKOLAS®

For fibers predicted by the AI model of MEKOLAS®, fiber length, fiber width, and aspect ratio were determined through subsequent image analysis. Fibers that met the fiber counting criteria of the Asbestos Monitoring Manual were selected as fibers to be counted. When fiber length and fiber width could not be accurately determined due to the shape of the detected fiber, it was classified as indeterminable.

A prediction was considered correct when the pixels of fibers designated as counting targets in the evaluation data overlapped with those of fibers predicted by MEKOLAS®, and those fibers were also identified as counting targets through image processing.

The accuracy of MEKOLAS® was evaluated using two metrics: detection accuracy and counting accuracy. Detection accuracy is determined solely by whether or not fibers are detected, without performing fiber counting. Overlapping fibers will be regarded as a single aggregate and any differences in the number of fibers will be disregarded. For counting accuracy, fibers to be counted are counted individually, even if they overlap.

Two metrics were used for accuracy: recall and precision. Recall is the ratio of fibrous materials identified as counting targets by an expert measurer that were correctly estimated by MEKOLAS®. Precision is the ratio of fibrous materials identified as counting targets by MEKOLAS® that were judged as such by an expert measurer. These are calculated using the following formulas:

Recall = true positive / (true positive + false negative) (1)

Precision = true positive / (true positive + false positive) (2)

where:

True positive: Among the fibrous materials estimated by MEKOLAS®, the number of fibers identified as counting targets by an expert measurer.

False negative: The number of fibers counted by the expert analyst that MEKOLAS® failed to detect (missed).

False positive: The number of fibers detected by MEKOLAS® that were not detected by the expert analyst (overcounted).

In addition, since recall and precision generally have a trade-off relationship, evaluation was also conducted using the F1 score. The F1 score is calculated as follows:

F1 score = 2 x recall x precision / (recall + precision) (3)

An example is shown in Fig. 1. The left side shows the detection and counting results by the expert analyst, while the right side shows those by MEKOLAS®. Red indicates fibrous materials selected as counting targets (fibers numbered 1–9 by the expert analyst), and blue indicates fibrous materials that were not counted because their size or aspect ratio did not meet the counting criteria. When evaluating by detection accuracy, two aggregates of overlapping fibers (Nos 2 and 3, and 6 and 7) are each counted as one without individual identification. In this case, the recall is 1.0 because all seven fiber aggregates are detected, and the precision is 0.88 because one of the eight detected aggregates is a false detection. When evaluating by counting accuracy, the two aggregates (Nos. 2, 3, and 6, 7) are counted as four individual fibers. Of the nine fibers to be counted, MEKOLAS® counted seven because it could not distinguish the overlapping fibers, resulting in a recall of 0.78. Of the eight fibers detected, one was a false detection, resulting in a precision of 0.88.

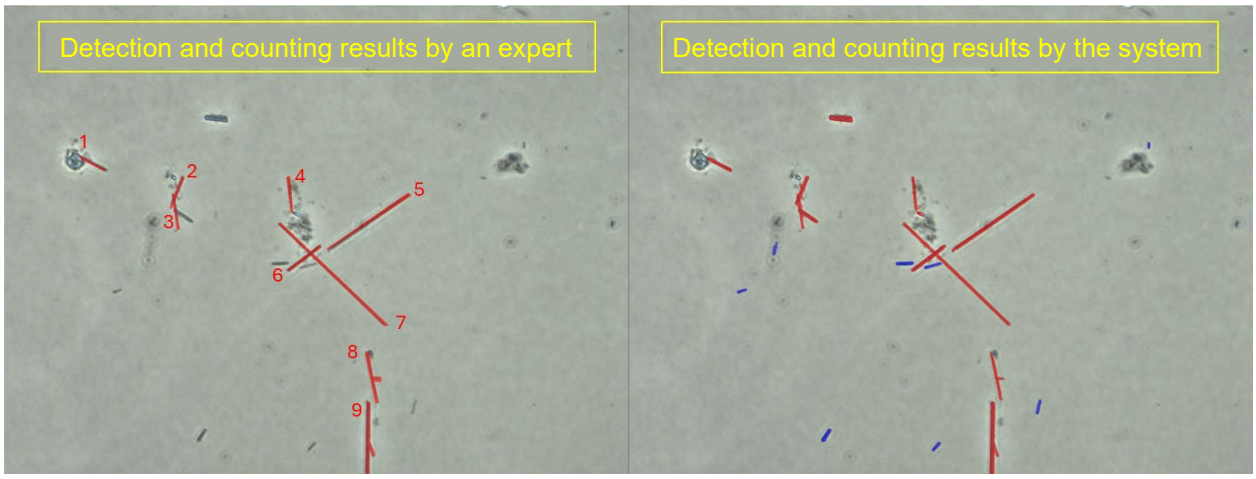

Fig. 1: Detection Accuracy and Counting Accuracy

2. Accuracy Evaluation Results

Tables 1 and 2 show the number of fibers and fiber aggregates detected by expert analysts, along with the fiber detection results of MEKOLAS®. When evaluated in terms of detection accuracy, the recall was 0.95, the precision was 0.71, and the F1 score was 0.81. There were 92 true positives (TP), 37 false positives (FP) representing false detections, and 5 false negatives (FN) representing missed detections. With few missed detections, the system detects fibers and their aggregates with high accuracy.

Tables 3 and 4 present the fiber counting results obtained by the expert analyst and the system. The number of fibers counted by the expert analyst per image ranged from 1 to 16, while the number of fibers detected by MEKOLAS® ranged from 1 to 14. The total fiber count by expert analysts across all images was 110. On the other hand, the system detected 133 fibers, consisting of 94 true positives (TP) and 38 false positives (FP). There were 16 false negatives (FN), indicating that false positives were more frequent than false negatives. With respect to fiber counting accuracy, MEKOLAS® achieved a recall of 0.85, a precision of 0.71, and an F1 score of 0.78.

Table 1: Expert Analyst Fiber Counts and MEKOLAS® Fiber Detection Results for PCM Images of Real Atmospheric Samples

Image	Expert analyst Count value	MEKOLAS®			
		True positive (TP)	False positive (FP)	False negative (FN)	Count value
A-7	1	1	1	0	2
A-8	1	1	0	0	1
A-12	1	1	0	0	1
A-16	1	1	0	0	1
B-3	1	1	2	0	3
B-4	4	3	1	1	3
B-11	1	1	0	0	1
B-13	3	3	0	0	3
C-2	3	3	2	0	5
C-6	11	11	4	0	17
C-11	7	7	0	0	7
D-4	1	1	5	0	6
D-7	5	5	2	0	6
D-10	7	7	1	0	8
D-12	2	2	1	0	4
D-14	6	6	1	0	7
E-6	15	14	3	1	18
E-7	7	6	4	1	11
E-15	11	10	4	1	16
E-16	4	3	3	1	6

E-19	5	5	3	0	8
	97	92	37	5	134

Table 2: Recall, Precision, and F1 Score for Each Image and for All Images in Detection Accuracy

Image	Recall	Precision	F1 Score
A-7	1	0.5	0.67
A-8	1	1	1
A-12	1	1	1
A-16	1	1	1
B-3	1	0.33	0.5
B-4	0.8	1	0.89
B-11	1	1	1
B-13	1	1	1
C-2	1	0.6	0.75
C-6	1	0.73	0.85
C-11	1	1	1
D-4	1	0.17	0.29
D-7	1	0.71	0.83
D-10	1	0.88	0.93
D-12	1	0.67	0.8
D-14	1	0.86	0.92
E-6	0.93	0.82	0.87
E-7	0.86	0.6	0.71
E-15	0.91	0.71	0.8
E-16	0.75	0.5	0.6
E-19	1	0.62	0.77
	0.95	0.71	0.81

Table 3: Fiber Counting Results by Expert Analysts and MEKOLAS® for PCM Images of Real Atmospheric Samples

Image	Expert analyst Count value	MEKOLAS®			
		True positive (TP)	False positive (FP)	False negative (FN)	Count value
A-7	1	1	1	0	2
A-8	1	1	0	0	1
A-12	1	1	0	0	1
A-16	1	1	0	0	1
B-3	1	1	2	0	3
B-4	5	3	0	2	4
B-11	1	1	0	0	1
B-13	3	3	0	0	3
C-2	3	3	2	0	5
C-6	12	12	4	0	16
C-11	6	6	1	0	7
D-4	1	1	5	0	6
D-7	7	4	2	3	6
D-10	9	7	1	2	8
D-12	4	3	1	1	4
D-14	8	6	1	2	7
E-6	16	14	4	2	18
E-7	9	7	4	2	11
E-15	12	11	4	1	15
E-16	4	3	3	1	6
E-19	5	5	3	0	8
	110	94	38	16	133

Table 4: Recall, Precision, and F1 Score for Each Image and All Images in Counting Accuracy

Image	Recall	Precision	F1 Score
A-7	1	0.5	0.67
A-8	1	1	1
A-12	1	1	1
A-16	1	1	1
B-3	1	0.33	0.5
B-4	0.57	1	0.73
B-11	1	1	1
B-13	1	1	1
C-2	1	0.6	0.75
C-6	0.81	0.76	0.79
C-11	1	1	1
D-4	1	0.17	0.29
D-7	0.57	0.67	0.62
D-10	0.78	0.88	0.82
D-12	0.75	0.75	0.75
D-14	0.75	0.86	0.8
E-6	0.83	0.83	0.83
E-7	0.78	0.64	0.7
E-15	0.8	0.75	0.77
E-16	0.75	0.5	0.6
E-19	0.83	0.62	0.71
	0.85	0.71	0.78

* This document consists of excerpts and re-edited content from the following papers:
Yukiko Iida, Takashi Yamamoto, Kazuharu Iwasaki, Ken-Ichi Yuki ,Kentaro Kiri, Hayato Yamashiro, Toshiyuki Toyoguchi, and Atsushi Terazono, Development of a rapid fiber-detection system using artificial intelligence in phase-contrast microscope images of actual atmospheric samples, Frontiers in Analytical Science, (25 June 2025), DOI : 10.3389/frans.2025.1571840, https://www.frontiersin.org/journals/analytical-science/articles/10.3389/frans.2025.1571840/full