Electronic Supplementary Information

Quantifying the CVD-Grown Two-Dimensional Materials via Image Clustering

Zebin Li[1,§], Jihea Lee[2,§], Fei Yao[2,*], Hongyue Sun[1,*]

[1]Department of Industrial and Systems Engineering, University at Buffalo, The State University of New York, Buffalo, NY, USA.

[2]Department of Material Design and Innovation, University at Buffalo, The State University of New York, Buffalo, NY, USA.

[§]These authors contributed equally

[*]Corresponding author: feiyao@buffalo.edu; hongyues@buffalo.edu

The detailed explanations or definitions of extracted features are shown below:

- *Solidity* is the measurement of the convex-concave degree of the shape, and it is calculated as the area of the flake over the convex area. In our case, if the solidity of the shape is close to 1, then the shape is close to a triangle.

- *Circularity* is the measurement of the roundness of the shape. It is calculated as $\dfrac{4 \times Area \times \pi}{Perimeter^2}$, and if the shape is close to a perfect circle, the value is close to 1.

- *Eccentricity* is the measurement of how close a conic section being a circular. In our case, it is calculated as the eccentricity of the ellipse that has the same second-moment of the shape. If the value is larger, then the shape is more narrow and close to a line segment.

- *Ratio of the equivdiameter over the perimeter*. The Euivdiameter of a shape is defined as the diameter of a circle that has the same area as the shape. The ratio of the equivdiameter over

the perimeter is a constant of 0.248 for an equilateral triangle. The difference from this constant can be used to measure the triangularity.

- *Ratio of the major axis length over the minor axis length*. The length of the major/minor axis of a shape is the length of the major/minor axis of the ellipse that has the same normalized second central moment as the shape, and it can be used as a measurement of the irregularity of MoS_2 flakes.

- *The number of MoS_2 flakes in each image*. This is a basic description of the optical images.

- *The number of light-spots*. Light-spots are the parts within MoS_2 flakes that have high brightness. They are multilayer MoS_2 formed during the growth and need to be avoided for monolayer samples.

- *Ratio of the number of light-spots over the number of MoS_2 flakes*. This is a feature derived from the above two features and can be used to measure the percentage of the multilayer flakes over monolayer ones.

- *Entropy* of each image, which is a statistical measurement of the randomness of images and is used to measure the image texture.

- *Perimeter of each MoS_2 flake*. This is a basic description of the MoS_2 flakes.

Table S1. CVD Synthesis Parameters

Images	Temperature (°C)	Dwell (min)	Flow rate (sccm)
Fig 5b – top left	740	40	100
Fig 5b – top right	760	50	100
Fig 5b – bottom left	780	10	100
Fig 5b – bottom right	780	20	100
Fig 5c – top left	800	40	100
Fig 5c – top right	780	20	100
Fig 5c – bottom left	780	40	100

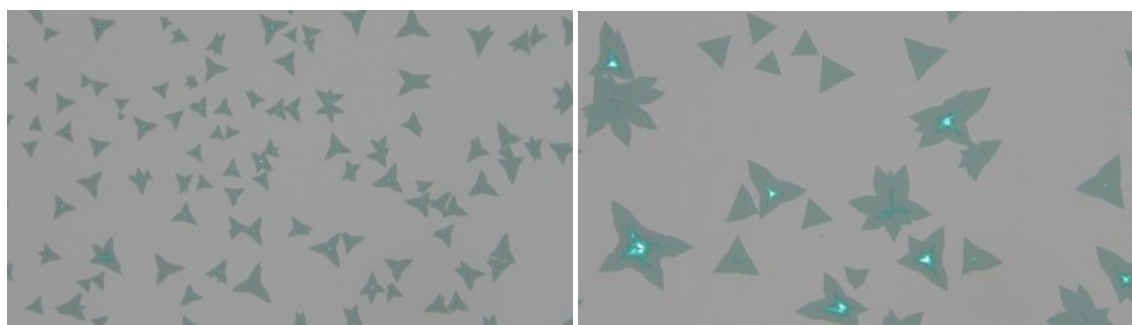|			
Fig 5c – bottom right	780	40	150
Fig 6a – top left	760	40	100
Fig 6a – top right	780	40	200
Fig 6a – bottom left	760	40	100
Fig 6a – bottom right	740	40	100
Fig 6b	780	10	100
Fig 7b – top left	800	40	100
Fig 7b – top right	780	40	200
Fig 7b – bottom left	780	20	100
Fig 7b – bottom right	780	20	100
Fig 7c	780	40	100

Fig. S1. The examples of optical images that are even hard for material experts to label.