

# CRISP: A deep learning architecture for GC×GC-TOFMS contour ROI identification, simulation, and analysis of imaging metabolomics

Vivek Bhakta Mathema<sup>1,2</sup>, Kassaporn Duangkumpha<sup>1,2</sup>, Kwanjeera Wanichthanarak<sup>1,2</sup>, Narumol Jariyasopit<sup>1,2</sup>, Esha Dhakal<sup>1,2</sup>, Nuankanya Sathirapongsasuti<sup>3,4</sup>, Chagriya Kitiyakara<sup>5</sup>, Yongyut Sirivatanauksorn<sup>2</sup>, Sakda Khoomrung<sup>1,2,5\*</sup>

<sup>1</sup>Metabolomics and Systems Biology, Department of Biochemistry, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

<sup>2</sup>Siriraj Metabolomics and Phenomics Center, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand

<sup>3</sup>Section of Translational Medicine, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand

<sup>4</sup>Research Network of NANOTEC - MU Ramathibodi on Nanomedicine, Bangkok, Thailand

<sup>5</sup>Department of Medicine, Faculty of Medicine, Ramathibodi Hospital, Rama VI Rd., Ratchathewi, Bangkok 10400, Thailand

<sup>6</sup>Center of Excellence for Innovation in Chemistry (PERCH-CIC), Faculty of Science, Mahidol University, Bangkok, Thailand

\*Corresponding author: [sakda.kho@mahidol.edu](mailto:sakda.kho@mahidol.edu)

**Source code, models and datasets** (Fully available only after acceptance of manuscript)

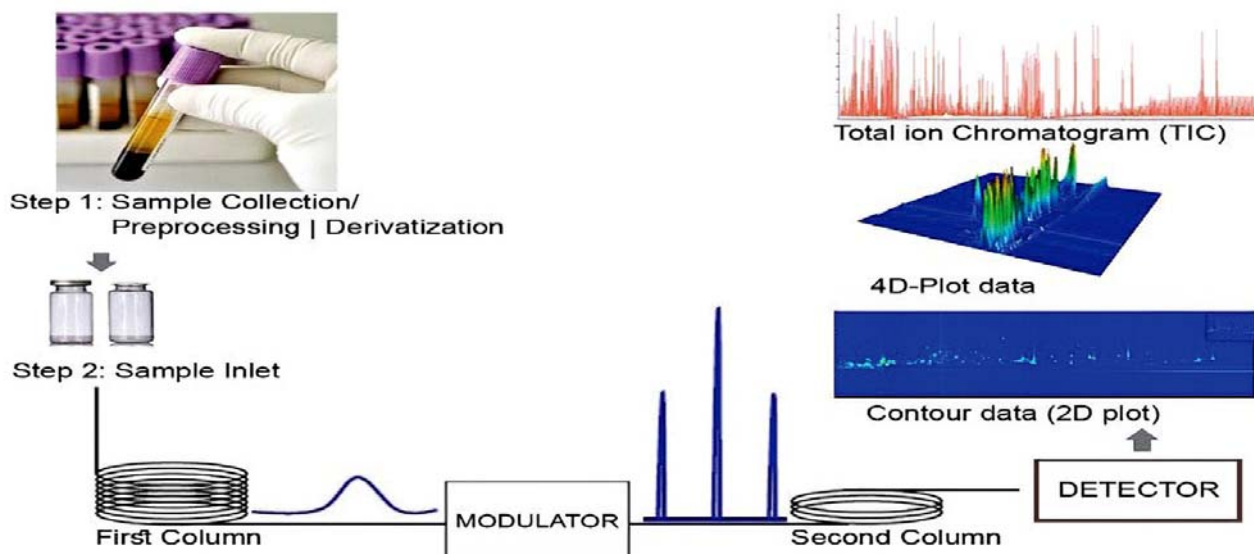
The source code and minimal example dataset, models are available in Github:

Weblink: <https://github.com/vivekmathema/GCxGC-CRISP>

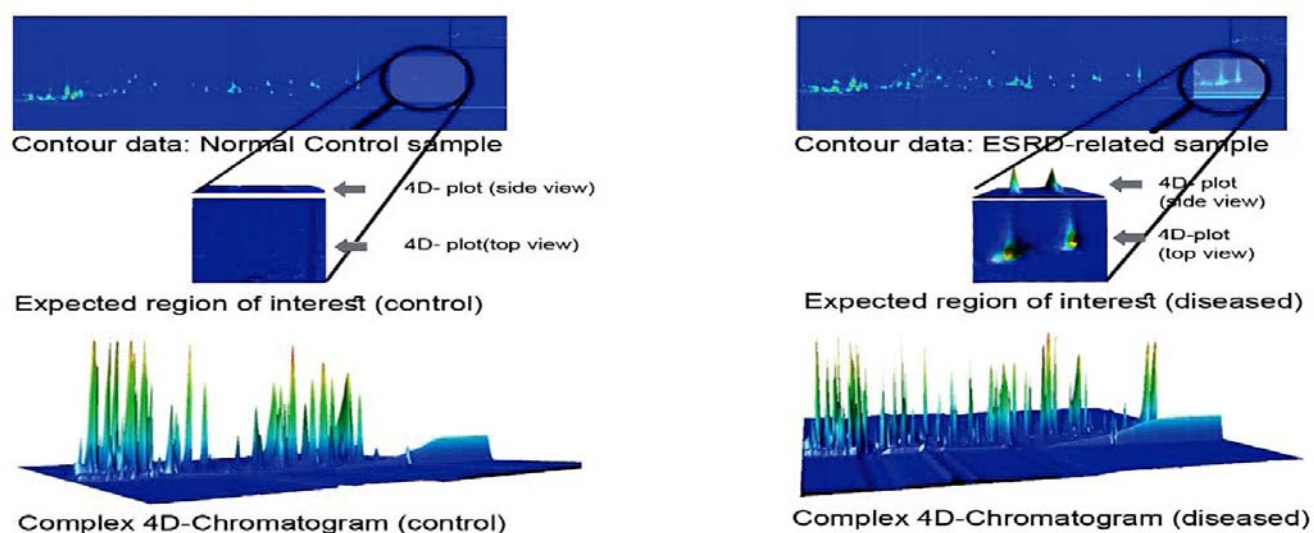
The manual for CRISP can be downloaded from Github:

Weblink: [https://github.com/vivekmathema/GCxGC-CRISP/blob/main/Manual/CRISP\\_manual.pdf](https://github.com/vivekmathema/GCxGC-CRISP/blob/main/Manual/CRISP_manual.pdf)

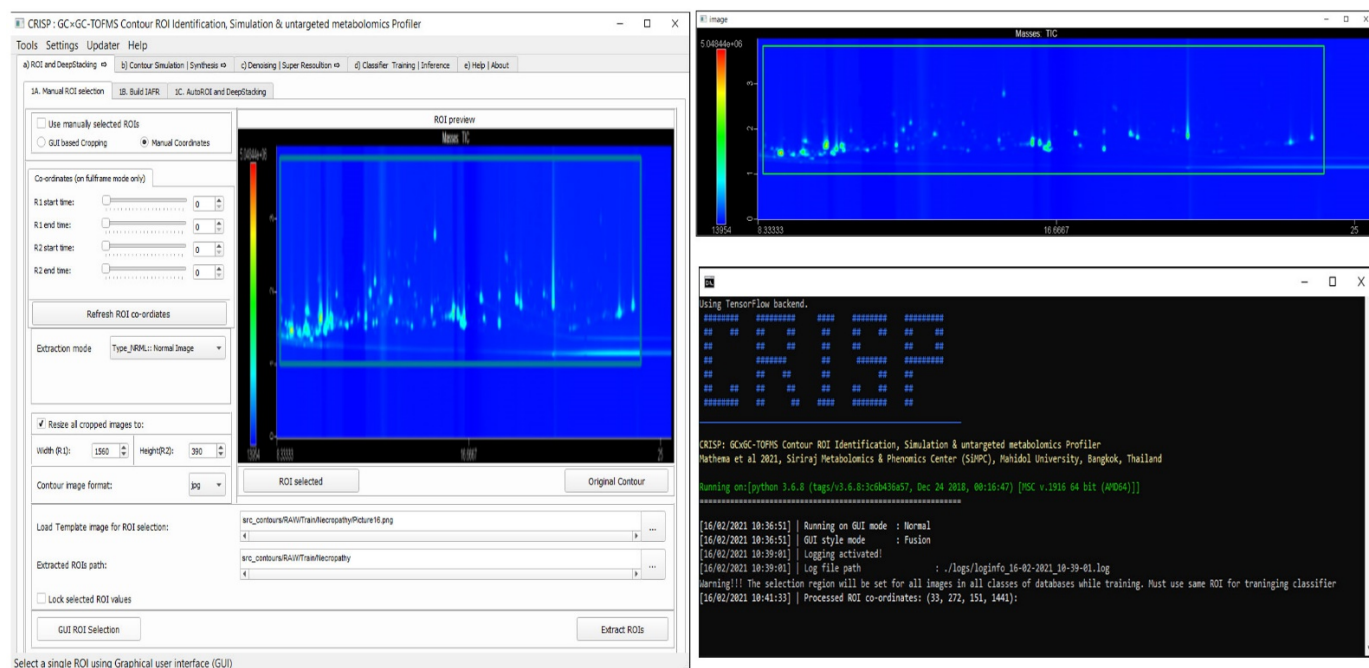
(A) GC×GC-TOFMS analysis of serum samples



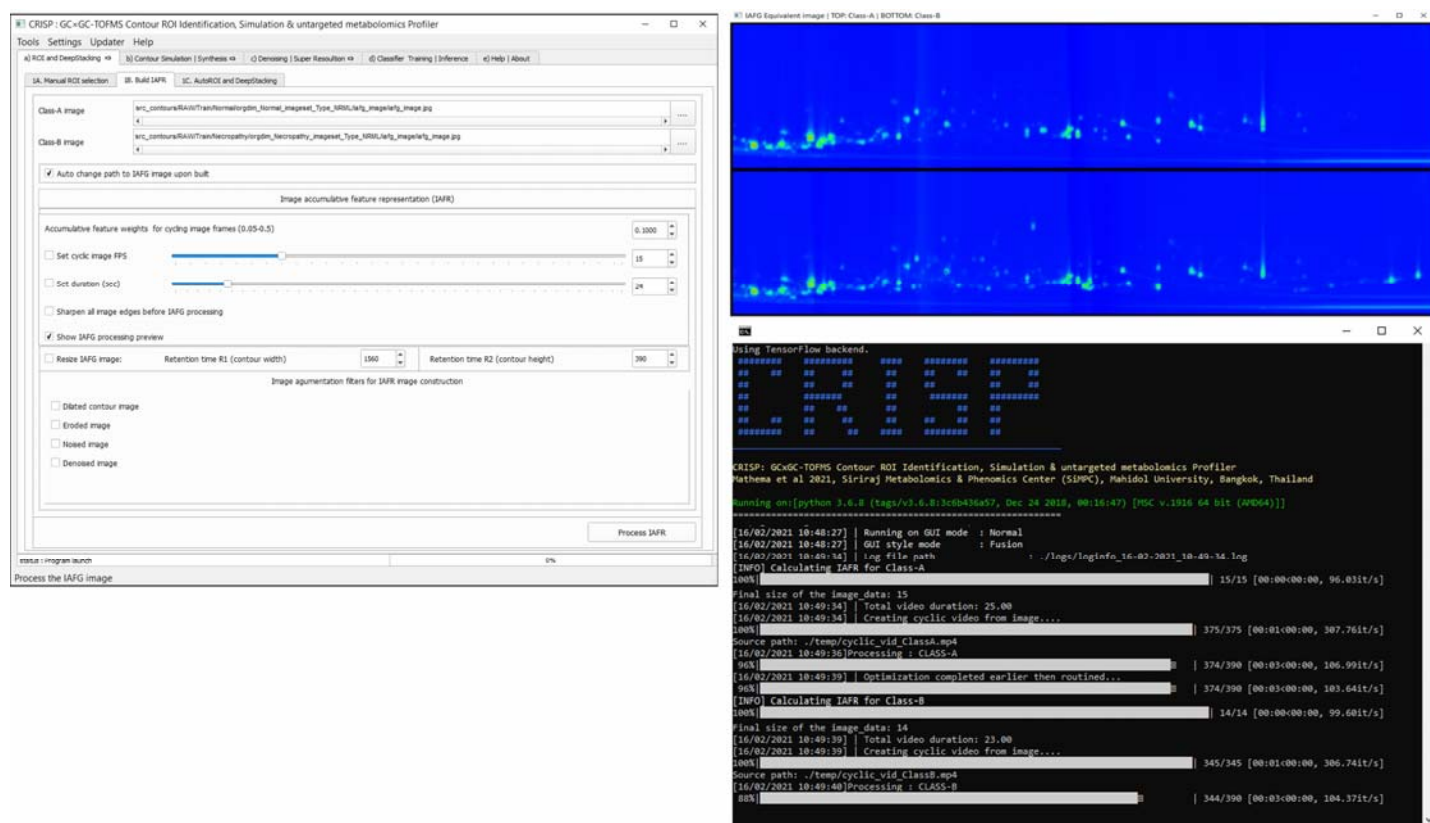
(B) Preliminary contrasting characteristics based on GC×GC-TOFMS



**Figure S1.** General schematics of the GC×GC-TOFMS indicating major steps involved in sample processing to generate contour data for healthy controls and the CKD/ESRD group. (A) Processing of the samples in GC×GC-TOFMS to obtain RAW as CDF file. (B) Automated analysis to obtain contour image data from ChromaTOF, showing some major differences between healthy and diseased group.



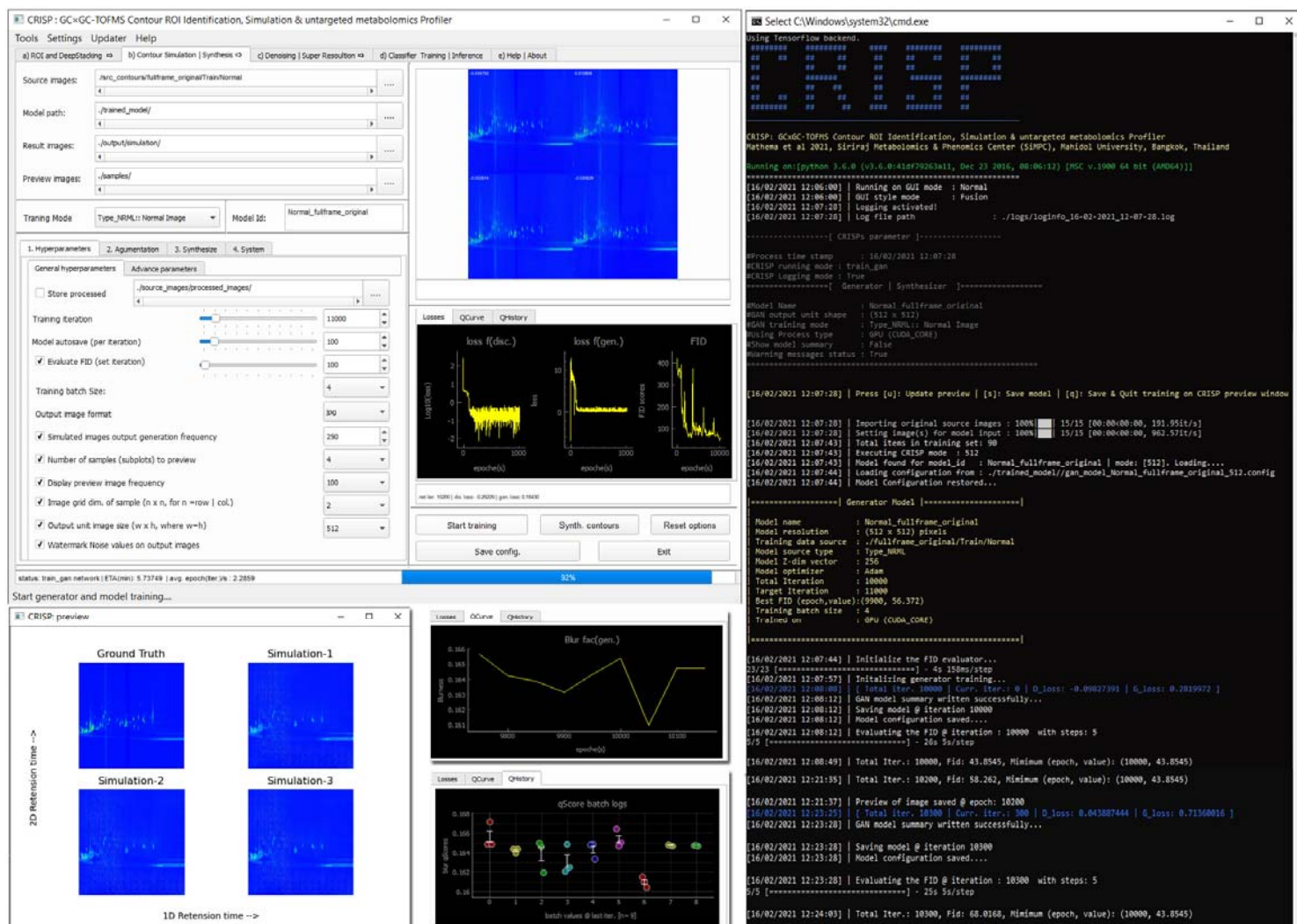
**Figure S2.** Single manual selection ROI selection. The screenshot of CRISP GUI and console screen showing example of a single manual ROI selection from whole GC×GC-TOFMS contour as indicated the boxed region in the RAW contour image.



**Figure S3.** AFRC construction from full frame GC×GC-TOFMS contour. Screenshot of CRISP module during construction of a single AFRC using default parameters for each of two groups (control vs ESRD/DM ).







**Figure S5.** Contour simulation and generator training. The screenshot of CRSIP module showing real-time generator training for GC×GC-TOFMS contour simulation. The graphs of generator, discriminator, and qScore are shown in real-time along with randomly selected real contour and user-defined number of simulated contours. The updates on console shows additional information of training progression including custom model summary.

```
C:\Windows\system32\cmd.exe
Using TensorFlow backend.
#####
CRISP: GCxGC-TOFMS Contour ROI Identification, Simulation & untargeted metabolomics Profiler
Mathema et al 2021, Siriraj Metabolomics & Phenomics Center (SIMP), Mahidol University, Bangkok, Thailand
Running on:[python 3.6.8 (v3.6.8:02f7028a11, Dec 23 2018, 08:06:12) [MSC v.1900 64 bit (AMD64)]]

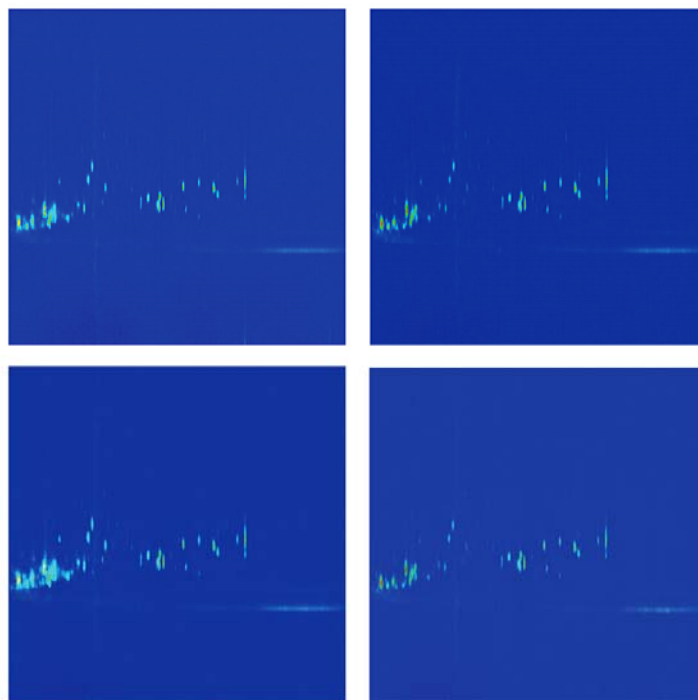
libomp warning: iccp: known incorrect simd profile
libomp warning: iccp: known incorrect simd profile
[16/02/2021 12:32:10] Running on GUI mode : Normal
[16/02/2021 12:32:10] GUI style mode : Fusion
[16/02/2021 12:33:13] Logging activated!
[16/02/2021 12:33:13] Log file path : ./logs/loginfo_16-02-2021_12-33-13.log
[16/02/2021 12:33:13] GPU Found & activated. Running -> [GPU MODE]

-----[ CRISP's parameter ]-----
#Process time stamp : 16/02/2021 12:33:13
#CRISP running mode : synthesis
#CRISP logging mode : True
-----[ Generator | Synthesizer ]-----
#Model Name : normal_fullframe_original
#MMH output unit shape : (512 x 512)
#MMH training mode : Type: MMH:: Normal Image
#Model path location : ./trained_model/
#Mixing Process type : GPU (CUDA_CORE)
#Show model summary : False
#Generating messages status: True

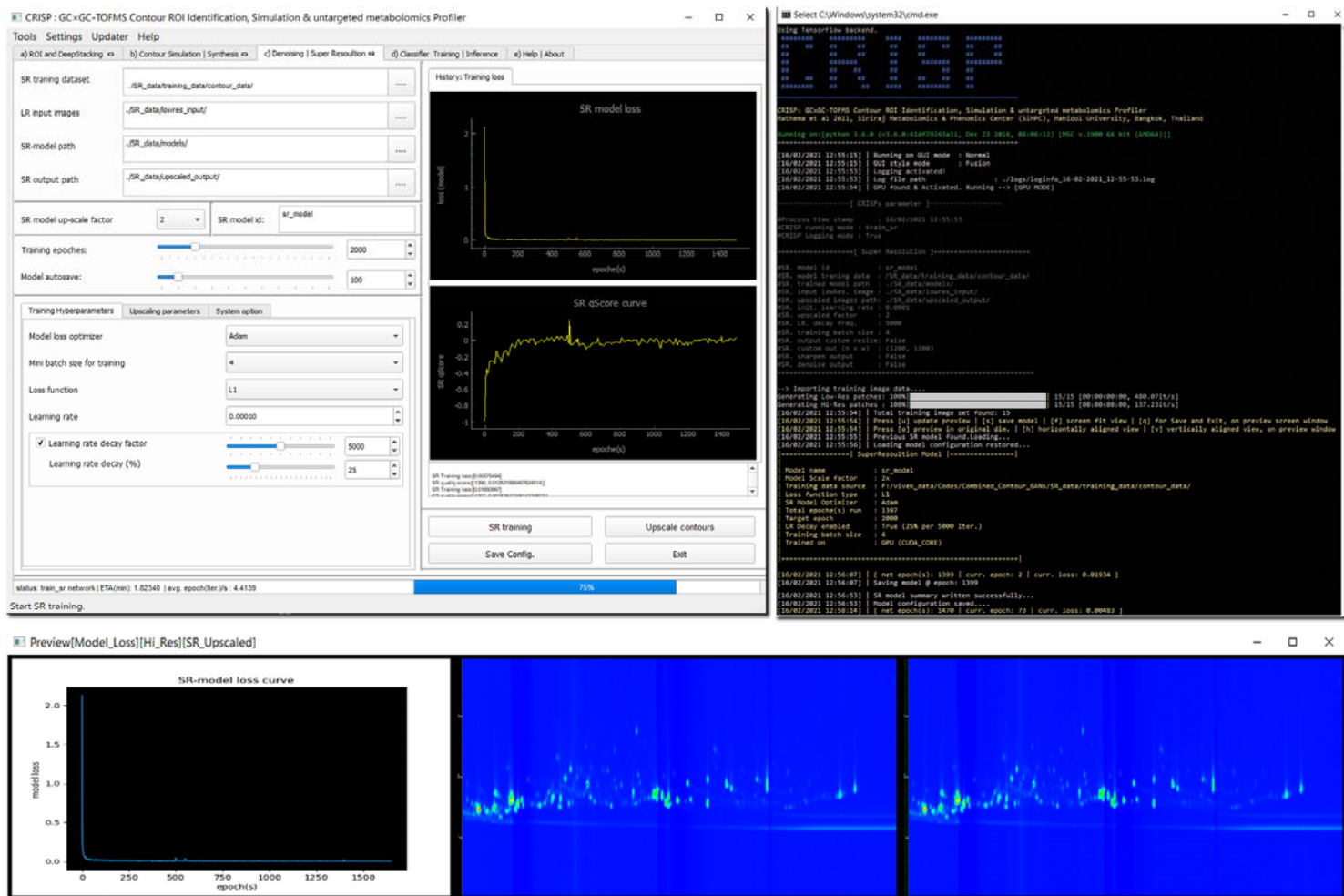
[16/02/2021 12:33:16] Loading trained weights for mode: [512]. Loading....
[16/02/2021 12:33:17] Model loaded successfully.....
[16/02/2021 12:33:17] Model configuration found.Loading model configuration...
[16/02/2021 12:33:17] Model Configuration restored...

-----[ Generator Model ]-----
Model name : normal_fullframe_original
Model resolution : (512 x 512) pixels
Training data source : ./fullframe_original/Train/Normal
Model source type : Type_MMH
Model 2-dim vector : 256
Model optimizer : Adam
Total Iteration : 100000
Target Iteration : 10000
Best F1D (epoch,value):(10000, 43.8545)
Training batch size : 4
Trained on : GPU (CUDA_CORE)

-----
[16/02/2021 12:33:17] Generating synthetic image: 100% [100/100 [00:19:00.00, 5.221it/s]]
[16/02/2021 12:33:37] Total of 100 contour(s) were successfully generated
[16/02/2021 12:33:37] Total of 0 augmented image(s) were added
```



**Figure S6.** Synthesis of GCxGC-TOFMS contours using trained model. On left, screenshot of CRISP console showing synthesis of user-defined number of contours. On right, images of four synthetic contours obtained using trained generator which can be used for classifier training. Model training history and customized summary of trained model is displayed during each model launch.



**Figure S7.** Superresolution model for contour quality improvement. The screenshot of the CRISP module ongoing training of GCxGC-TOFMS contour super resolution model for improvement of contour quality. The GUI shows real-time model loss along with qScore and preview of real as well as resolution-enhanced (2x) image during training session. The updates on console also shows detail information of training progression including custom model summary.

```
C:\Windows\system32\cmd.exe
Using TensorFlow backend.
*****
**  **  **  **  **  **
**  **  **  **  **  **
**  **  **  **  **  **
**  **  **  **  **  **
**  **  **  **  **  **
**  **  **  **  **  **
*****

CRISP: GCxGC-TOFMS Contour ROI Identification, Simulation & untargeted metabolomics Profiler
Mathema et al 2021, Siriraj Metabolomics & Phenomics Center (SIMP), Mahidol University, Bangkok, Thailand

Running on:[python 3.6.0 (v3.6.0:41df79263a11, Dec 23 2016, 08:06:12) [MSC v.1900 64 bit (AMD64)]]

[16/02/2021 13:25:04] Running on GUI mode : Normal
[16/02/2021 13:25:04] GUI style mode : fusion
[16/02/2021 13:26:34] Logging activated!
[16/02/2021 13:26:34] Log file path : ./logs/loginfo_16-02-2021_13-26-34.log
[16/02/2021 13:26:36] GPU found & Activated. Running --> [GPU MODE]

-----[ CRISPs parameter ]-----

Process time stamp : 16/02/2021 13:26:35
CRISP running mode : SR Upscaling
CRISP Logging mode : True

-----[ Super Resolution ]-----

SR model id : sr_model
SR model training data : ./SR_data/training_data/DIV2K/
SR trained model path : ./SR_data/models/
SR input lowRes image : ./SR_data/lowres_input/
SR upscaled images path : ./SR_data/upscaled_output/
SR init learning rate : 0.0001
SR upsample factor : 2

[16/02/2021 13:32:46] Loading SR model...
[16/02/2021 13:32:47] Previous SR model found.Loading...
[16/02/2021 13:32:48] Loading model configuration restored...
[16/02/2021 13:32:48] Model Configuration restored...

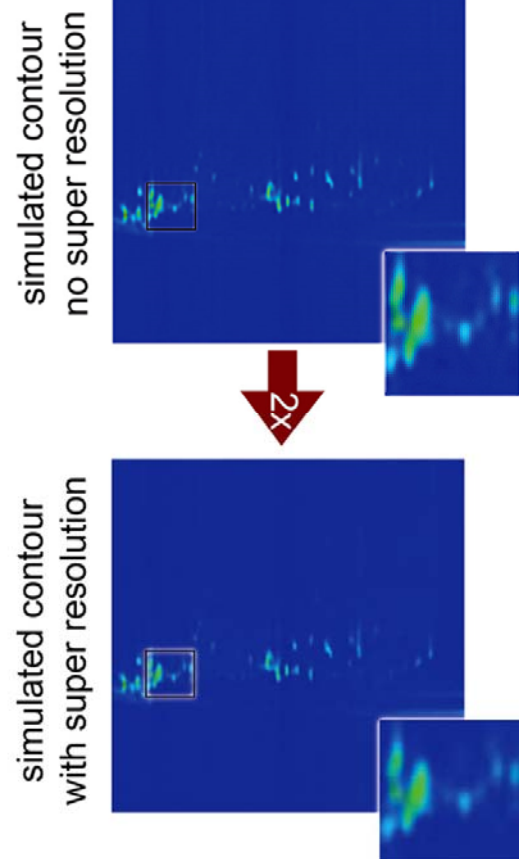
-----[ SuperResolution Model ]-----

Model name : sr_model
Model Scale factor : 2x
Training data source : ./SR_data/training_data/DIV2K/
Loss function type : L1
SR Model Optimizer : Adam
Total epoche(s) run : 1649
Target epoch : 1000
LR Decay enabled : False
Training batch size : 4
Trained on : GPU (CUDA_CORE)

-----

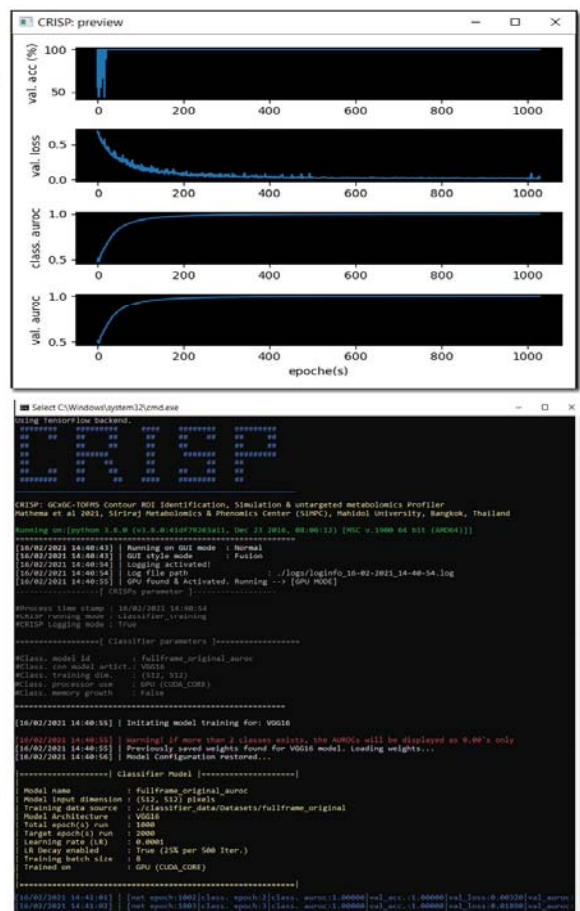
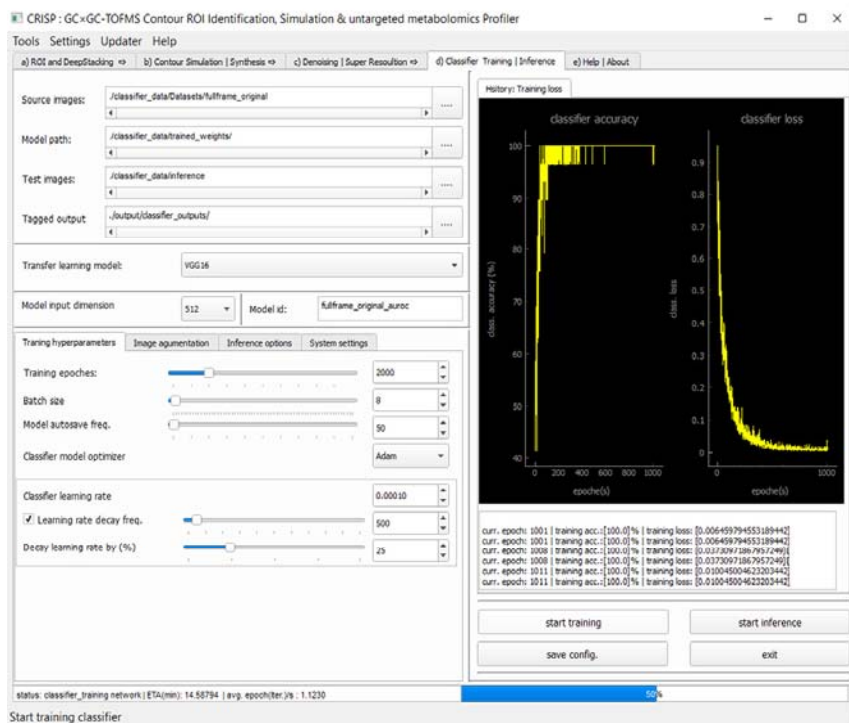
[16/02/2021 13:32:48] Total images to upscale : 5
-> Processing image: ./SR_data/lowres_input/simulated_512_0.jpg
-> Processing image: ./SR_data/lowres_input/simulated_512_12.jpg
-> Processing image: ./SR_data/lowres_input/simulated_512_13.jpg
-> Processing image: ./SR_data/lowres_input/simulated_512_14.jpg
-> Processing image: ./SR_data/lowres_input/simulated_512_15.jpg

[16/02/2021 13:32:50] Total of 5 images up_scaled succesfully...
```



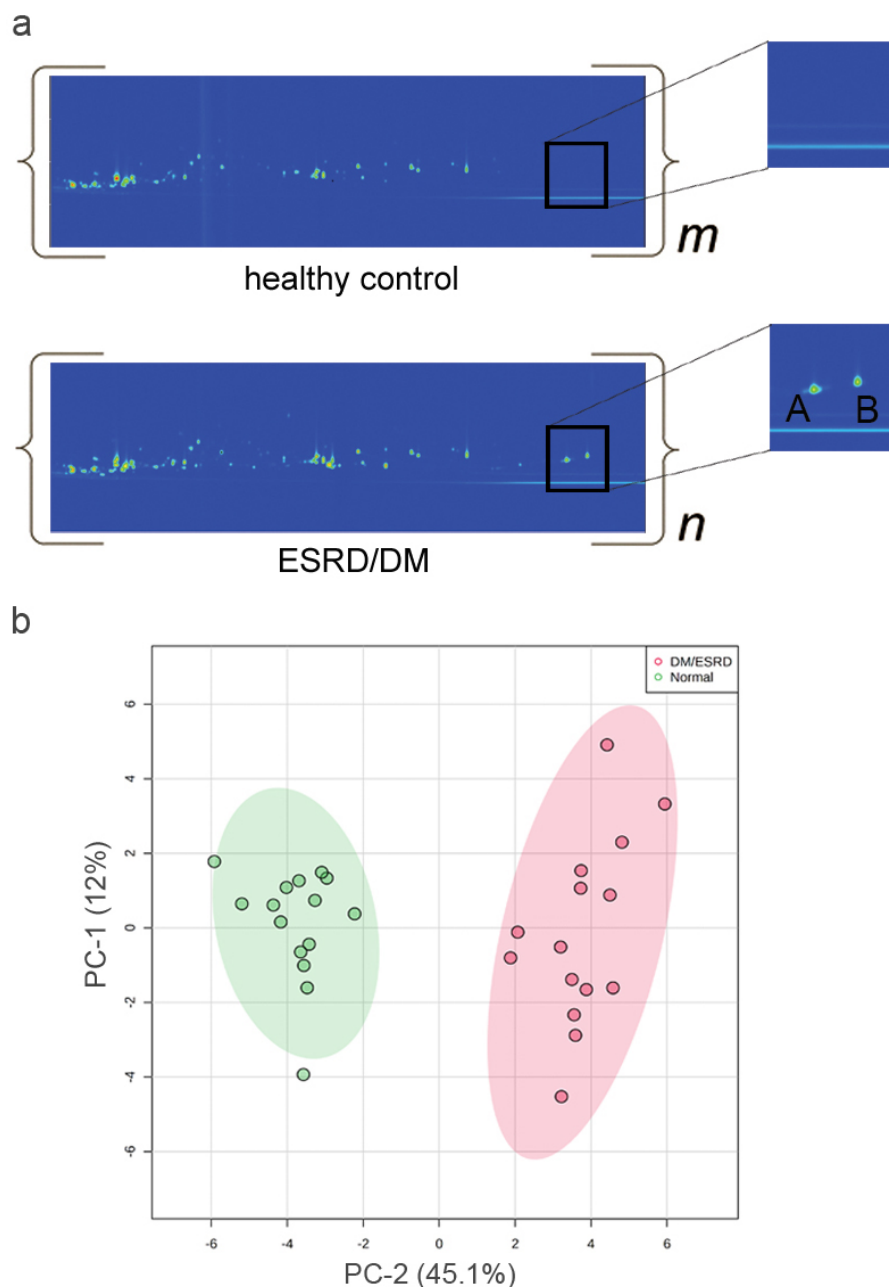
**Figure S8.** Super resolution enhancement of GCxGC-TOFMS simulated contour images. The screenshot showing ongoing image quality enhancement of simulated contour images based on the custom trained super resolution model. The left panel shows console information containing model summary and progression of the resolution enhancement process. The right panel shows a zoomed example of the improvement on simulated GCxGC-TOFMS contour resolutions based on custom trained CRISP super resolution model. The updates on console shows detailed information of trained model including custom model summary.





**Figure S9.** Classifier training using transfer learning. Screenshot of CRISP module for ongoing classifier training based on transfer learning method using VGG16 architecture. Real-time update graphs for model classification and validation accuracies, model losses and AUROC curves. The console screen displays the model summary and ongoing model training progress.





**Figure S11.** Contrasting feature between ESRD/DM and healthy control GC×GC-TOFMS contour feature. a) The contrasting features between GC×GC-TOFMS contour image for and healthy control (top) and ESRD/DM (bottom). The metabolite A and B represents maltitol and luctitol, respectively b) The PCA results for ESRD/DM and healthy control showing difference in metabolite profile based on GC×GC-TOFMS raw data. ESRD/DM, End-Stage Renal Disease with diabetes mellitus.