Artificial intelligence (AI) is all over the place. DigiM’s AI engine might not be as flashy as self-driving car or data mining of consumer’s internet behavior. We are tackling a simple, yet notorious problem that haunts every image analyst – automatic image segmentation.

**Challenge Statement**

Imaging data are expressed as greyscale values. We will discuss in a separate article regarding how, and how much quantitative information can be extracted directly from greyscale value. For most analysis workflows, though, image quantification and image-based simulation, rely on segmented images.

Image segmentation is a critical step in image-based analytics. It transforms greyscale values into material phase identifications. People might not be always aware of this step. For example, we use simple measurement tools to estimate the size of a feature from an X-Ray MicroCT image. The exact feature and how the images are collected are not important for this context, but it happens to be a crystalline drug particle, as shown in Figure 1. When we measure it, our brain implicitly defines the boundary of the particle (blue contour in the insert). This is, conceptually, image segmentation.

For both quantitative image analytics algorithms (e.g., computing porosity and pore distribution) and image-based simulations (e.g., permeability), segmentation results are needed as the inputs. Consequently, it is critical to get accurate segmentation to avoid garbage-in garbage-out situation. In another word, without a critical review on the segmentation, it is difficult to believe any numbers computed.
Figure 1. Implicit segmentation during a linear measurement of a feature in an image

The dominant segmentation algorithms, available commercially or as open source, can be categorized as,

a. Greyscale intensity value based algorithm such as thresholding and its variations.

b. Gradient of greyscale value based algorithms, such as local threshold, top hat, watershed, etc.

In our example image, Figure 1, four material phases can be identified from the sample, as shown in Figure 2 left. While the bright phase, crystalline API, and the darkest phase, pores, can both be segmented with thresholding reasonably well, it is very difficult to segment amorphous API from pure polymer excipient matrix, due to their inter-mingled intensity value. Figure 2 right shows an attempt with low threshold. In polymer domain, a lot of pixels were missing using this threshold value, while a lot of pixels in amorphous drug domain are already segmented incorrectly as polymer.
Typical Segmentation Challenge Example

Low threshold: insufficient and noisy segmentation

Figure 2. Material phase identification (left), image histogram (upper right) and segmentation attempt on Polymer Excipient phase with low threshold (lower right).

If we continue to increase the threshold, we got the result in Figure 3 right. Amorphous API and polymer phases are completely mixed and segmented out as one phase. This example illustrates that it is impossible to find the right threshold that can segment out amorphous API from polymer.
**Typical Segmentation Challenge Example**

High threshold: over-segmentation

*Intensity and intensity gradient based methods fail in most cases.*

![Image showing material phase identification](image_url)

**Figure 3.** Material phase identification (left, same as Figure 2 left), image histogram (upper right) and segmentation attempt on Polymer Excipient phase with high threshold (lower right).

Similar to this example, the above-mentioned greyscale threshold or gradient algorithms seldom works on real-world images. The image analysts are left with three options,

1. Spend tremendous amount of effort in image filtering to enhance the contrast to the level that threshold or gradient method will work.
2. Resort to manual segmentation (which is extremely time consuming for large 3D dataset).
3. Blame the microscopists.

**DigiM’s AI Segmentation Solution**

When our eyes can recognize a feature, they use not only the greyscale value of pixels, but also its relationship with its surrounding pixels. The collection of the pixels reflects a unique signature of a material in response to the imaging signal, as a textural pattern.
DigiM’s artificial intelligence engine first learns these patterns from human, and populates that knowledge to all the images. Figure 4 illustrates a typical workflow using the web interface of DigiM I2S Enterprise Data Management System.

Figure 4. DigiM AI image segmentation workflow.
The segmentation results corresponding to Figure 2 and Figure 3 are shown in Figure 5. All four phases are identified with reasonable accuracy. Figure 6 shows a 3D visualization of polymer excipient phase. For those with curious minds, the four phases are quantified to be,

- Porosity, 2.62%
- Polymer excipient, 40.13%
- Amorphous API, 47.17%
- Crystalline API, 62%

Particle and pore size distributions are plotted in Figure 7, an extract from DigiM I2S.

DigiM AI (Artificial Intelligence)
Segmentation Example

*Figure 5. Material phase identification (left, same as Figure 2 left), image histogram (upper right) and segmentation attempt on Polymer Excipient phase with DigiM AI (lower right).*
Figure 6. 3D volume visualization of polymer excipient phase.

Figure 7. 3D particle size distribution of all four phases.
Concluding remarks

It is remarkable that for some even simple, two-phase systems, how much DigiM’s AI segmentation can help to tremendously reduce both the need of aggressive filtering, and uncertainty. AI is not perfect, just as human intelligence. There are times that we have to work on the imaging side to reduce damaging artifacts, which is precisely why DigiM also works intensively on imaging (either internally or with our clients).

There is actually one advantage of AI on human intelligence. Regardless how much I love image processing, and how many images I’ve processed with my talented DigiM colleagues, I guarantee that I will make very different mistakes in the evening after a long day, from those I made in the same morning, even though they maybe equally dumb. At least, when AI makes a mistake, it will make that mistake consistently.

DigiM’s AI capability is available through DigiM research contract, DigiM I2S subscription or academic collaboration. Contact us at info@digimsolution.com.

Acknowledgement

Data courtesy of Joe Neilly and Gerald Danzer from AbbVie, Inc is gratefully acknowledged. None of DigiM’s work is possible without the inspiration and ground work by the Open Source community – thank you all!