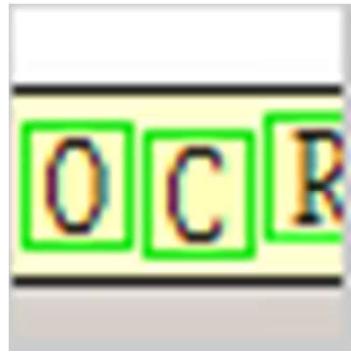
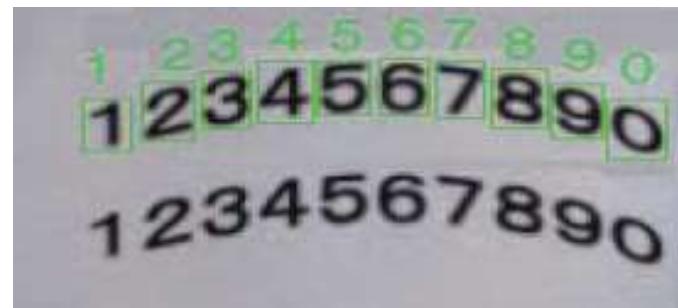


# Thuật toán OCR



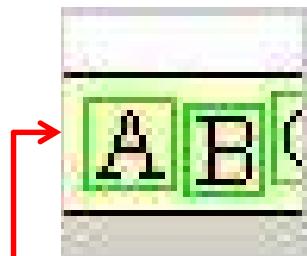
x. Optical Character Recognition  
(Nhận dạng ký tự quang học)

Nhận dạng các ký tự bằng  
các từ điển có sẵn



**OMRON**

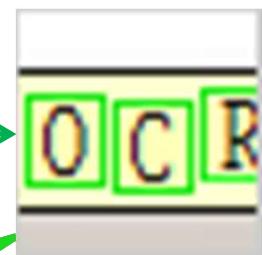
# OCR – Các công cụ có hiệu lực



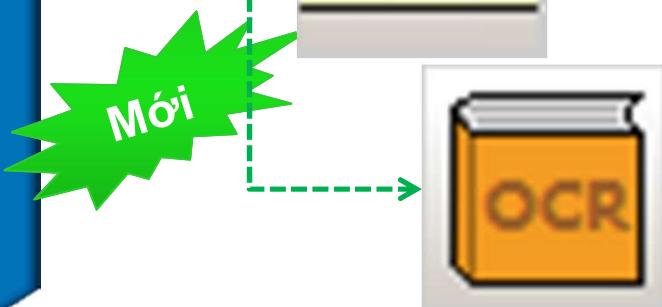
Nhận dạng lý tự



Tùy điển mẫu (**Bắt buộc**)



OCR – Nhận dạng ký tự quang học



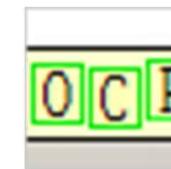
OCR Tùy điển người dùng (**Tùy chọn**)

OMRON

# Character Inspection vs. OCR



Character  
Inspection



OCR

Methods	segmentation	Edge based	- Line Extraction - Optimized Character Segmentation
	recognition	Template matching	Statistical approaches
Characteristic	[Pro]easy to adjust parameters [Con]must set up the dictionary [Con]speed increase in proportion to the registered characters	[Pro]Built-in dictionary (easy to combined add-on dictionary) [Pro]robust to touching characters, curved strings [Pro]supported to extract to a maximum of 4 lines	
Applications	FZ, FH(Ver1) *Character Inspection	FH(Ver2) *OCR	



OMRON

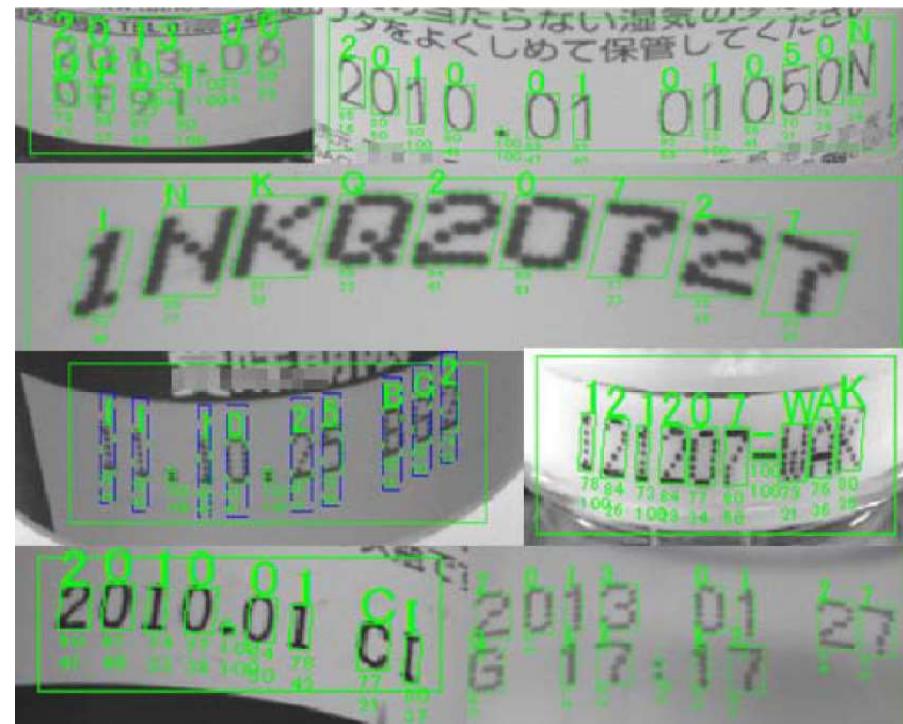
# OCR – Độ ổn định 1

Tăng cường độ ổn định cho việc nhận dạng ký tự:

## Các ký tự kiểu phẳng



## Các mảng ký tự bị cong



# OCR Độ ồn định 2 (Video)

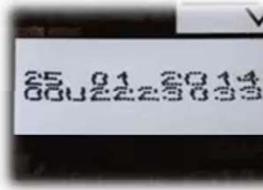


Think & See **OCR** *the OMRON Character Recognition*

**OMRON**

**Touching characters are often inevitable due to recent extremely fast line speed**

   
**Touching of characters**

   
**Touching lines      Touching with barcodes**

**Touching characters**

SYSMAC  
always in control

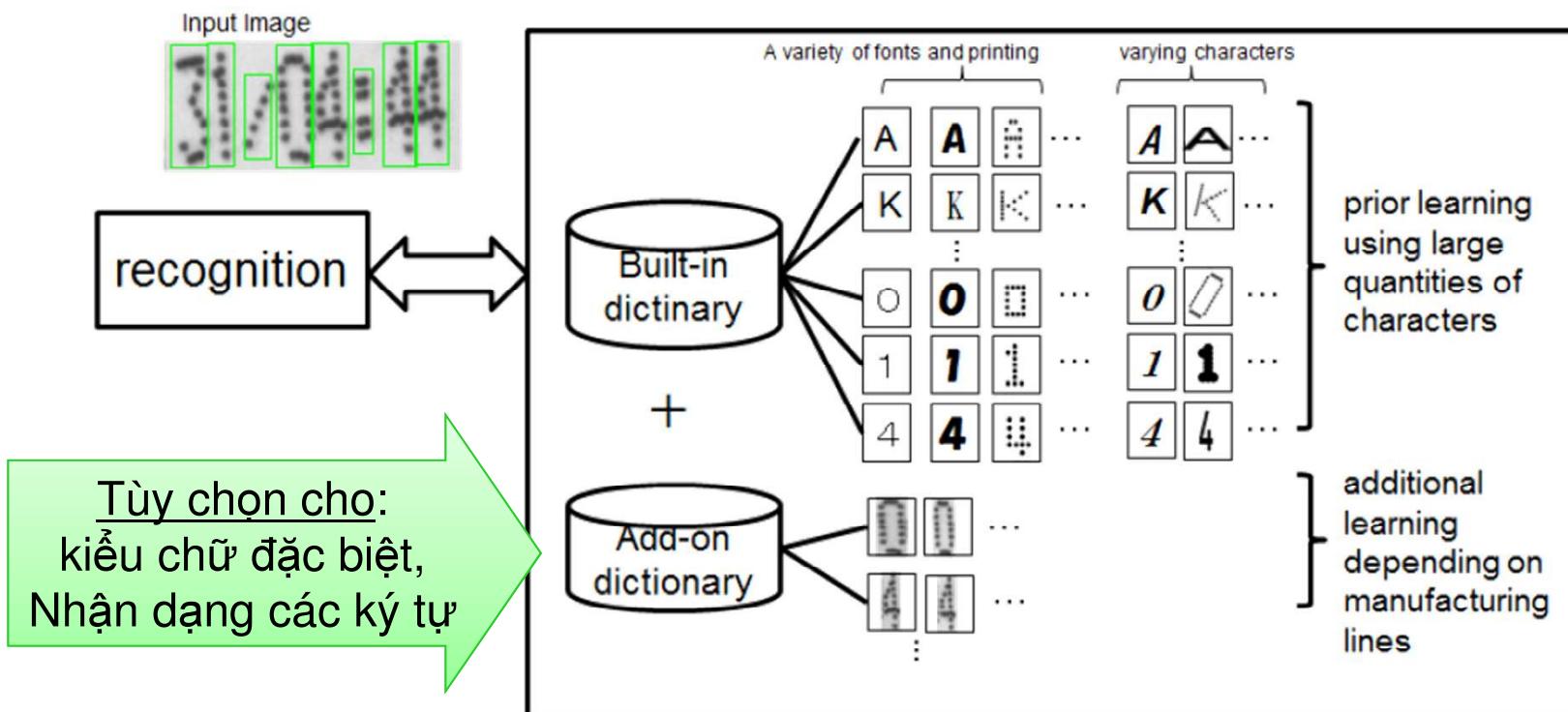
©OMRON Corporation 2014 All Rights Reserved

**OMRON**

# OCR – Khái niệm xây dựng từ điển

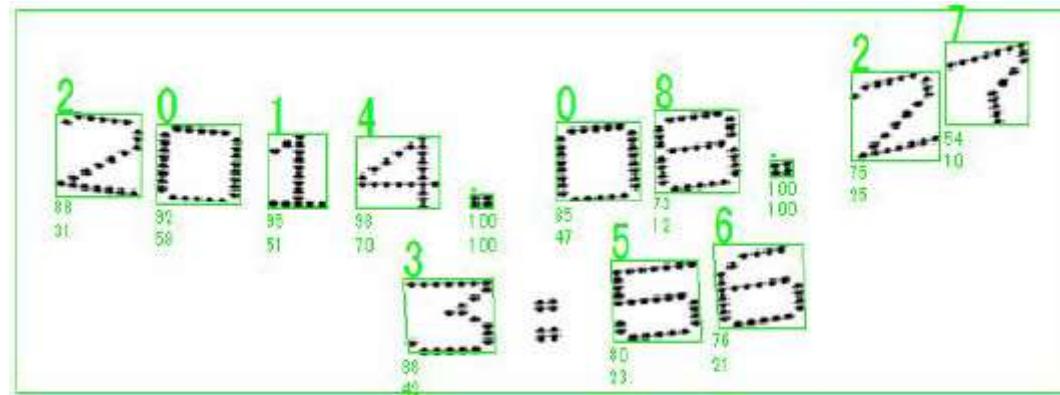
Nguồn dữ liệu ảnh ký tự lớn :

- Hơn 80 kiểu chữ được sử dụng
- Cộng thêm nhận dạng các kiểu chữ in: kích thước ký tự, nền, độ biến dạng và độ bóng



# OCR – Khái niệm sự phân đoạn

Các mảng sẽ tách ra và chia thành các đoạn ký tự

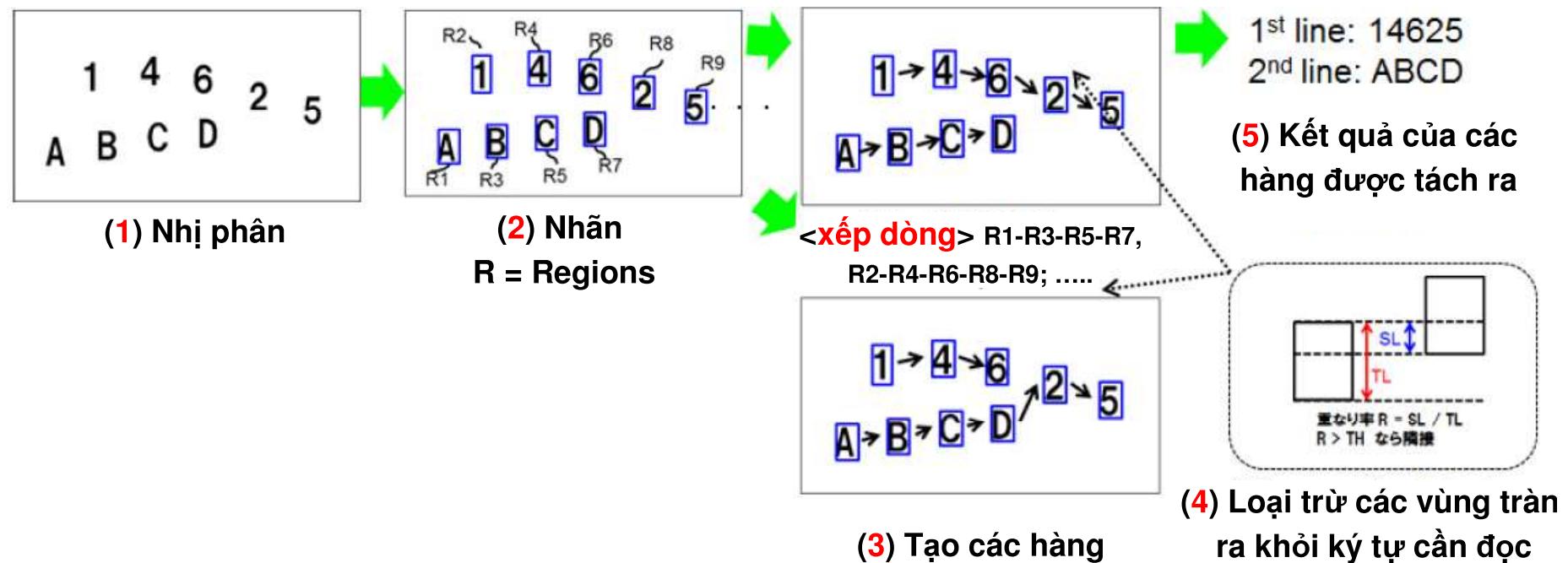
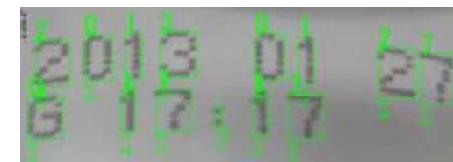


Có 2 kiểu phân đoạn thường được sử dụng:

- Line Extraction (Tách dòng)
- Optimized Character Segmentation (Tối ưu sự phân đoạn ký tự)

# OCR – Kiểu tách dòng ký tự

- Chủ yếu cho các ký tự cong!
- Nếu nhận dạng không được, bạn có thể kết hợp và/hoặc để đọc lại các thông tin quan trọng như:
  - Đọc ngày tháng trước
  - Số đếm sản xuất, ngày, giờ

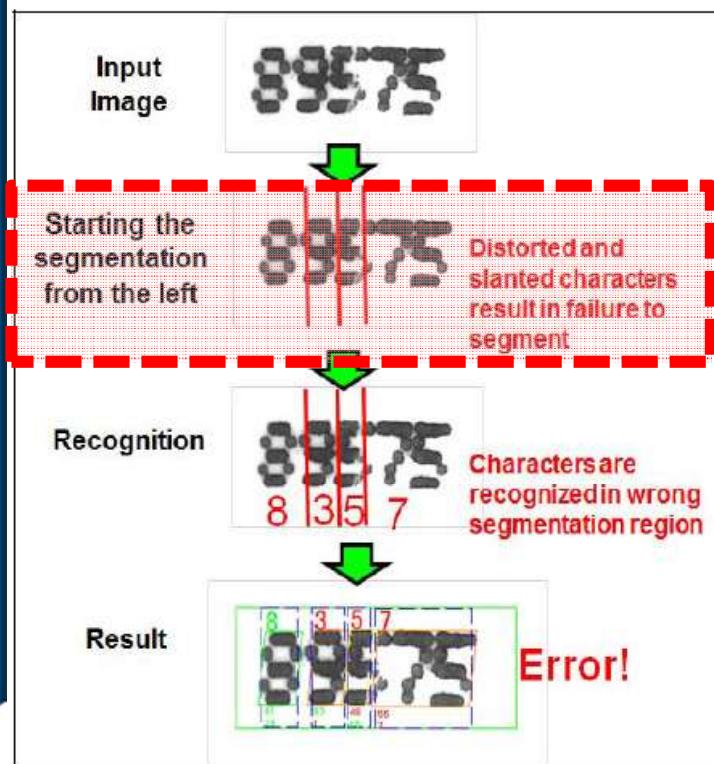


# OCR – Kiểu tối ưu phân đoạn ký tự

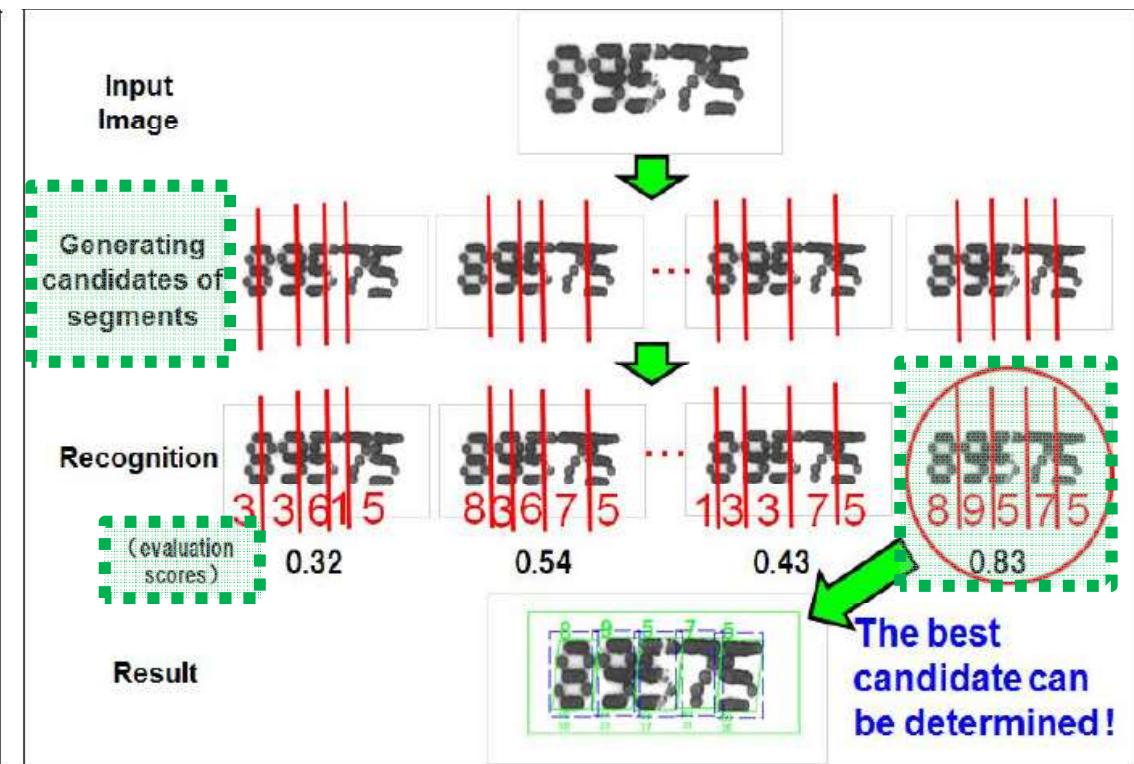


- Tự tạo các đoạn ký tự
- Nhận dạng và làm chúng tự ưu tiên hơn

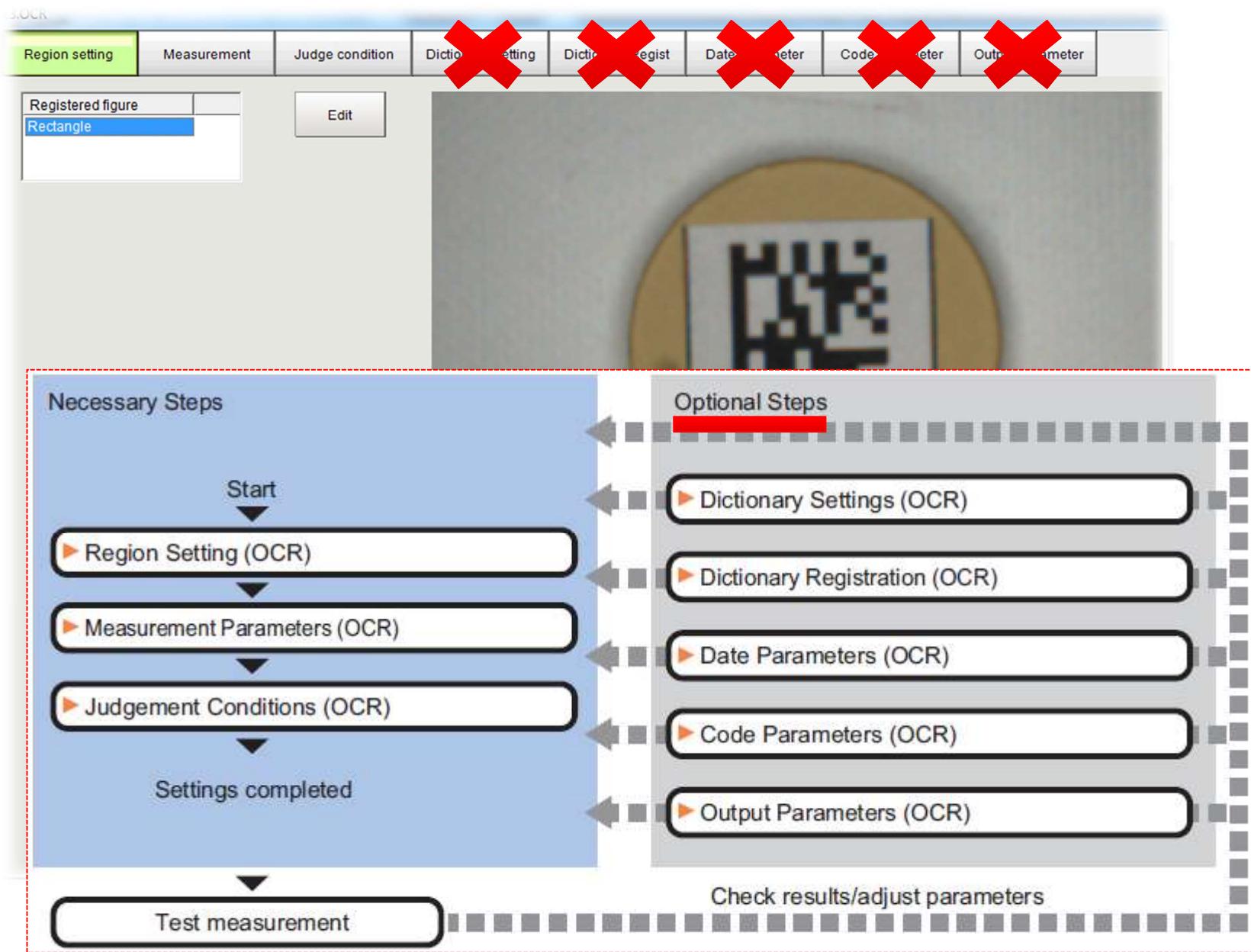
Kiểu nhận dạng trước đây



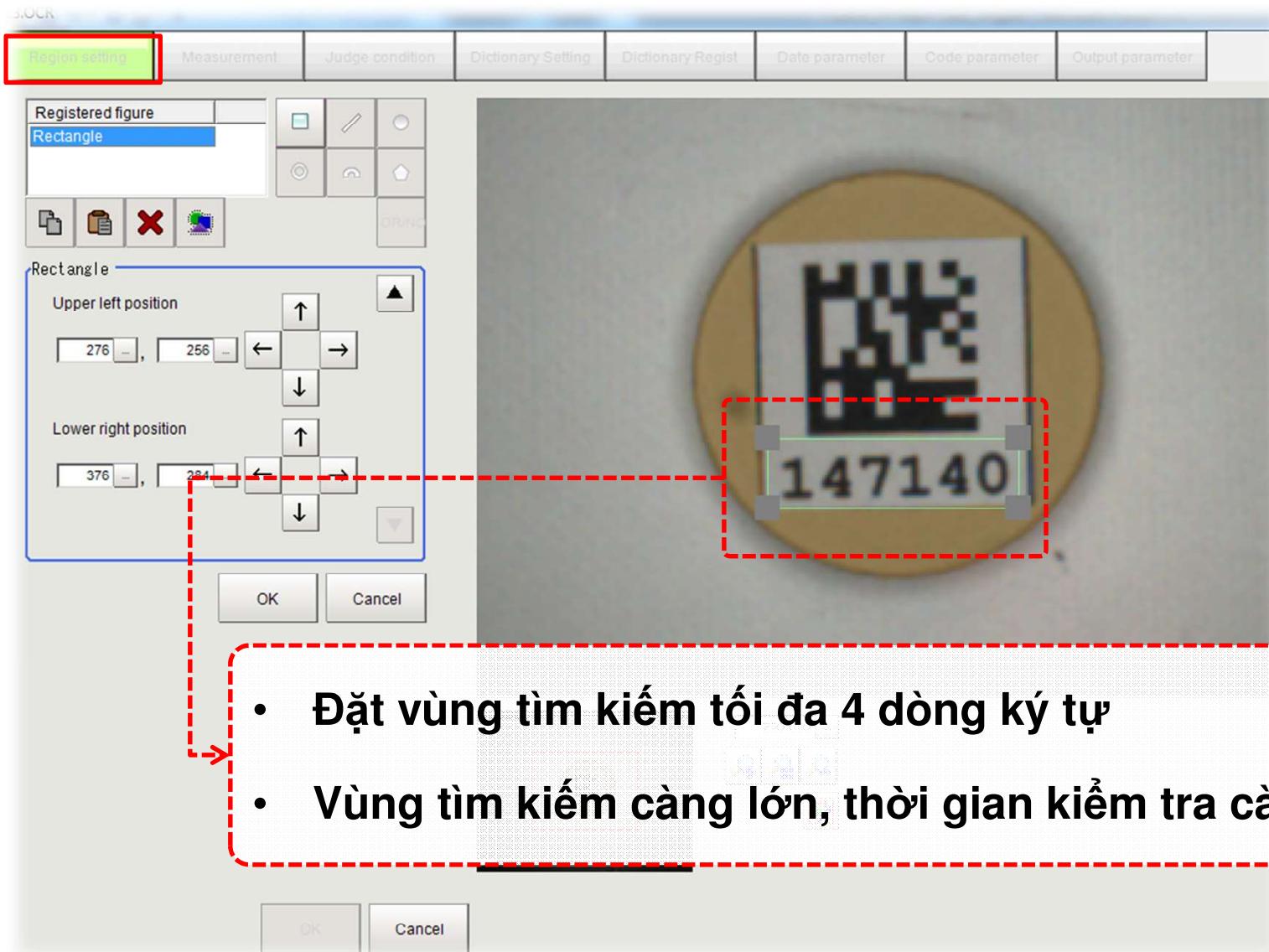
Kiểu nhận dạng OCR



# OCR – Cần thiết vs. Tùy chọn



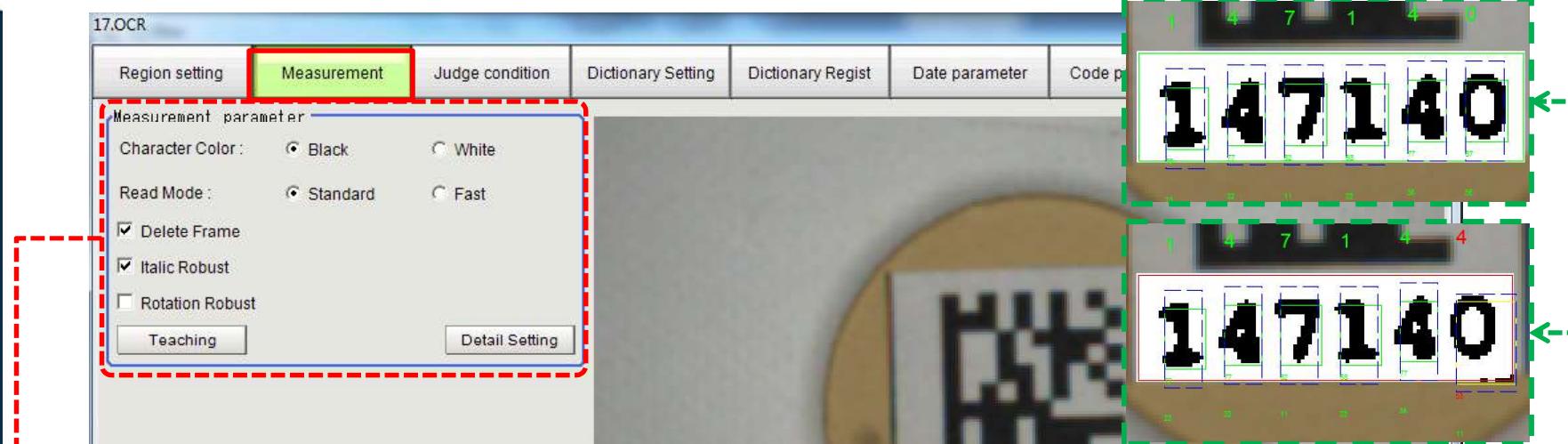
# OCR – Tab vùng tìm kiếm



- Đặt vùng tìm kiếm tối đa 4 dòng ký tự
- Vùng tìm kiếm càng lớn, thời gian kiểm tra càng lâu

**OMRON**

# OCR – Tab đo lường



- **Màu sắc ký tự** : đen/trắng
- **Chế độ đọc** : Kiểu chuẩn / nhanh (khi các ký tự cách xa nhau ta đặt là ‘fast’)
- **Xóa khung** : Xóa bỏ các điểm ảnh bị nhiễu đen tại vùng cần kiểm
- **Kiểu chữ nghiêng** : Cho phép nhận dạng ký tự nghiêng
- **Kiểu chữ xoay** : Cho phép khi các ký tự phải xoay

# OCR – Dạy học 1



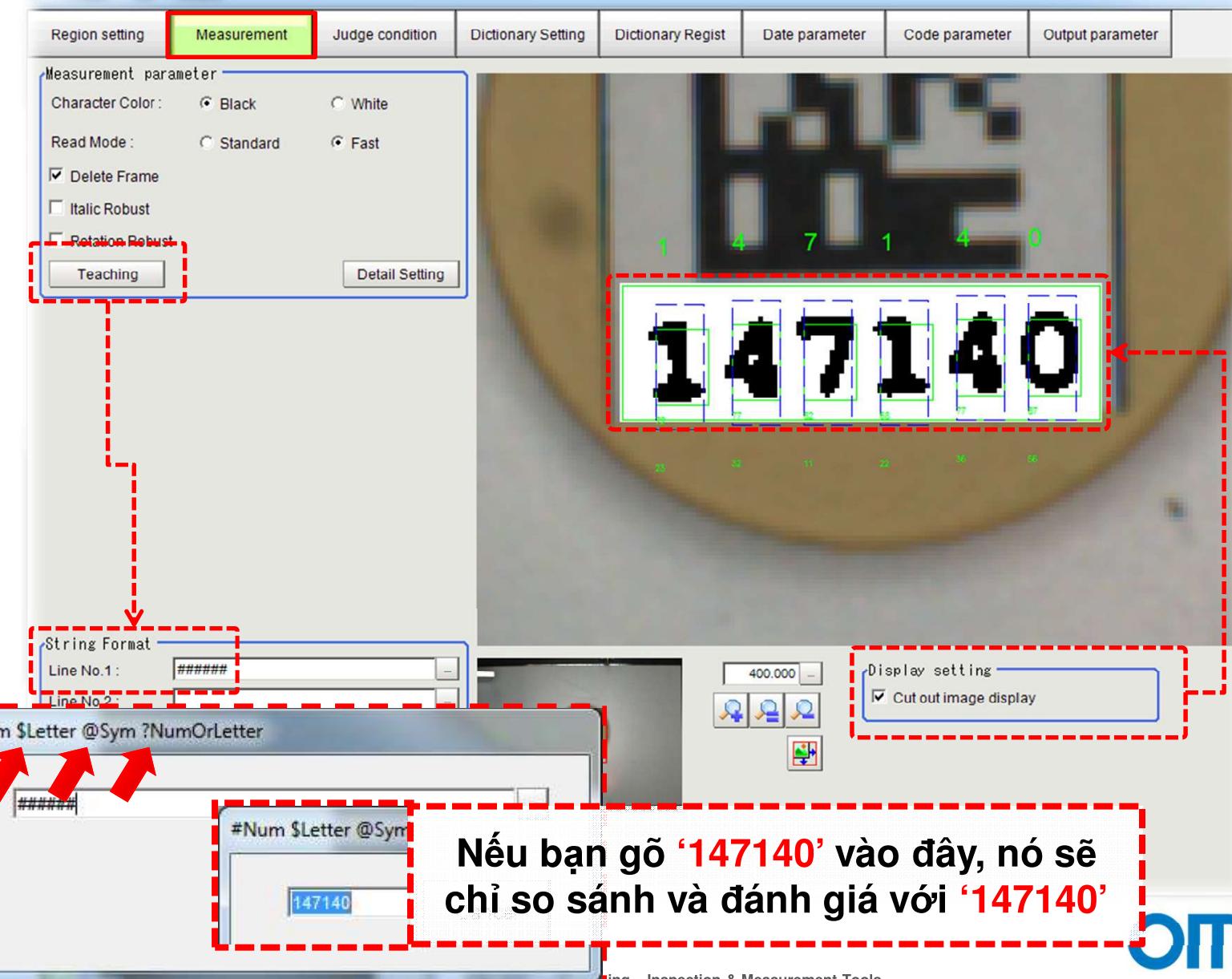
The screenshot illustrates the OCR teaching process across three windows:

- Main Window (Left):** Shows the "Measurement" tab selected. It includes settings for Character Color (Black), Read Mode (Standard), and various checkboxes like Delete Frame, Italic Robust, and Rotation Robust. A "Teaching" button is highlighted with a red dashed box and arrow.
- Teaching Window 1 (Top Right):** A dialog box titled "Teaching" with a red border. It contains a checkbox "Teach without correct result" and four input fields for "Correct String (Line No.1)" through "Line No.4". The first field contains "OR".
- Teaching Window 2 (Middle Right):** A dialog box titled "Teaching" with a green border. It contains a checkbox "Teach without correct result" and four input fields for "Correct String (Line No.1)" through "Line No.4". The first field contains "147140". A green arrow points to this field, and a green stamp with the word "RECOMMENDED" is visible on the right.
- Teaching Window 3 (Bottom Right):** A dialog box titled "Teaching" with a red border. It contains a checkbox "Teach without correct result" and four input fields for "Correct String (Line No.1)" through "Line No.4". The first field contains "#####".

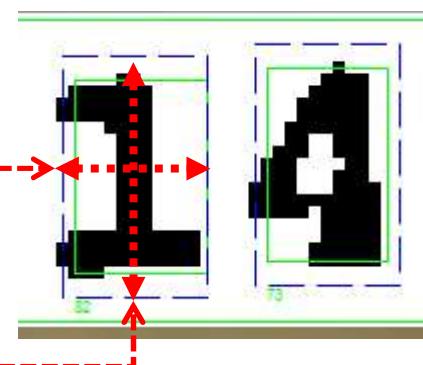
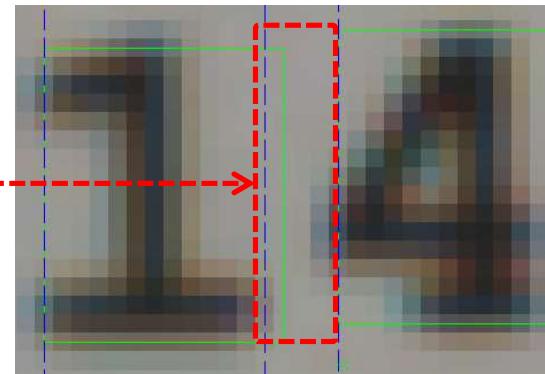
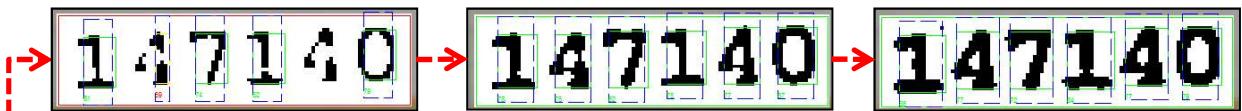
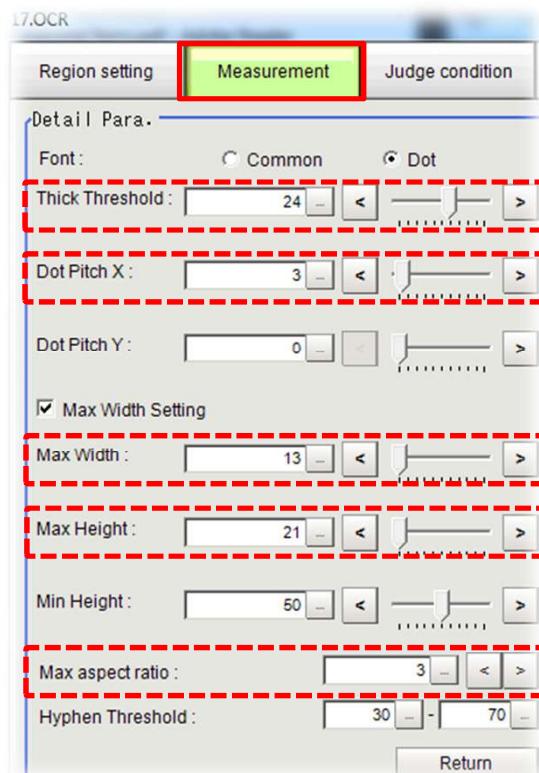
Below the main window, there are sections for "String Format" (with Line No.1-4 input fields) and "Display setting" (with a scale bar).

# OCR Unit – Dạy học 2

L7.OCR



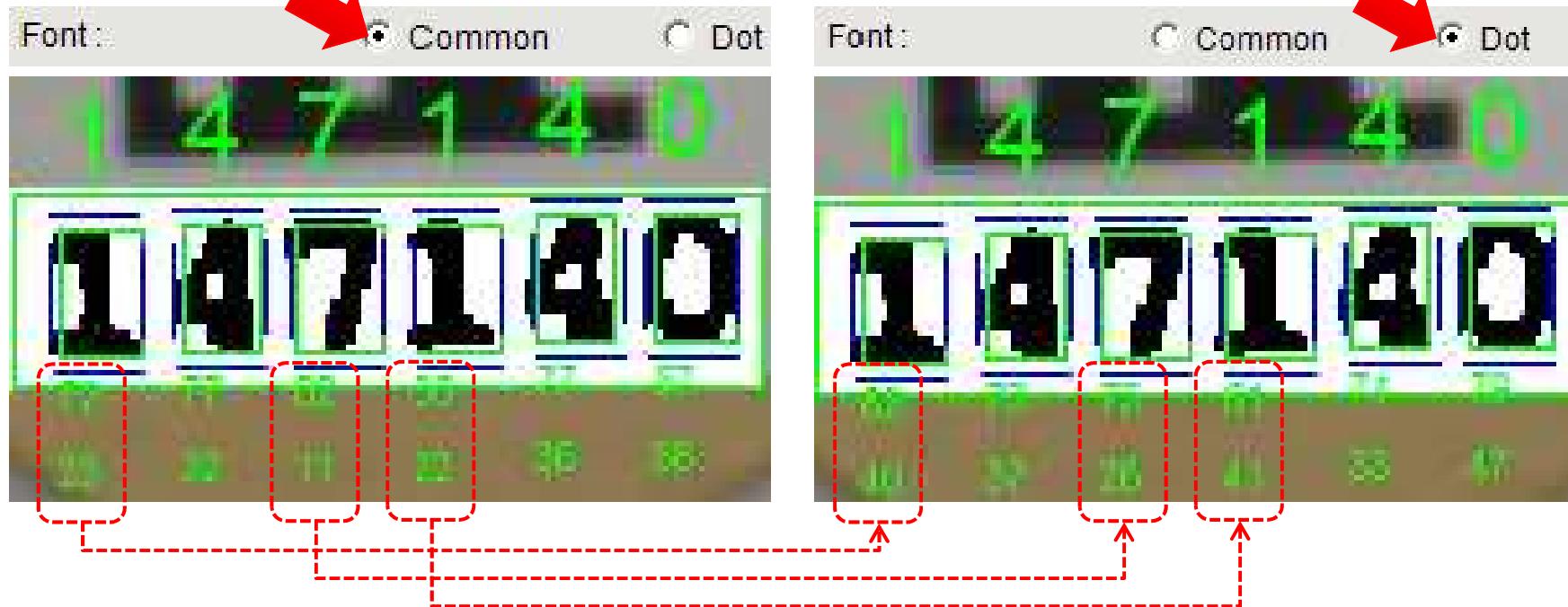
# OCR – Dạy học (Cài đặt chi tiết)



Nếu chiều cao/rộng > '3' → nhận dạng như 1  
kiểu ký tự hoặc ký hiệu đặc biệt

**OMRON**

# OCR – Cài đặt và kiểm tra?



Luôn cần theo dõi các tham số trên vì:

**Sự đồng dạng [60]:** Mức giống nhau giữa ký tự và thông tin kiểu chữ (được xây dựng trong từ điển)

**Sự ổn định [10]:** khác biệt giữa sự đồng dạng thứ nhất và sự đồng dạng thứ hai



**Giá trị thấp** = Cần thay đổi để đạt giá trị cao hơn tránh lỗi xảy ra



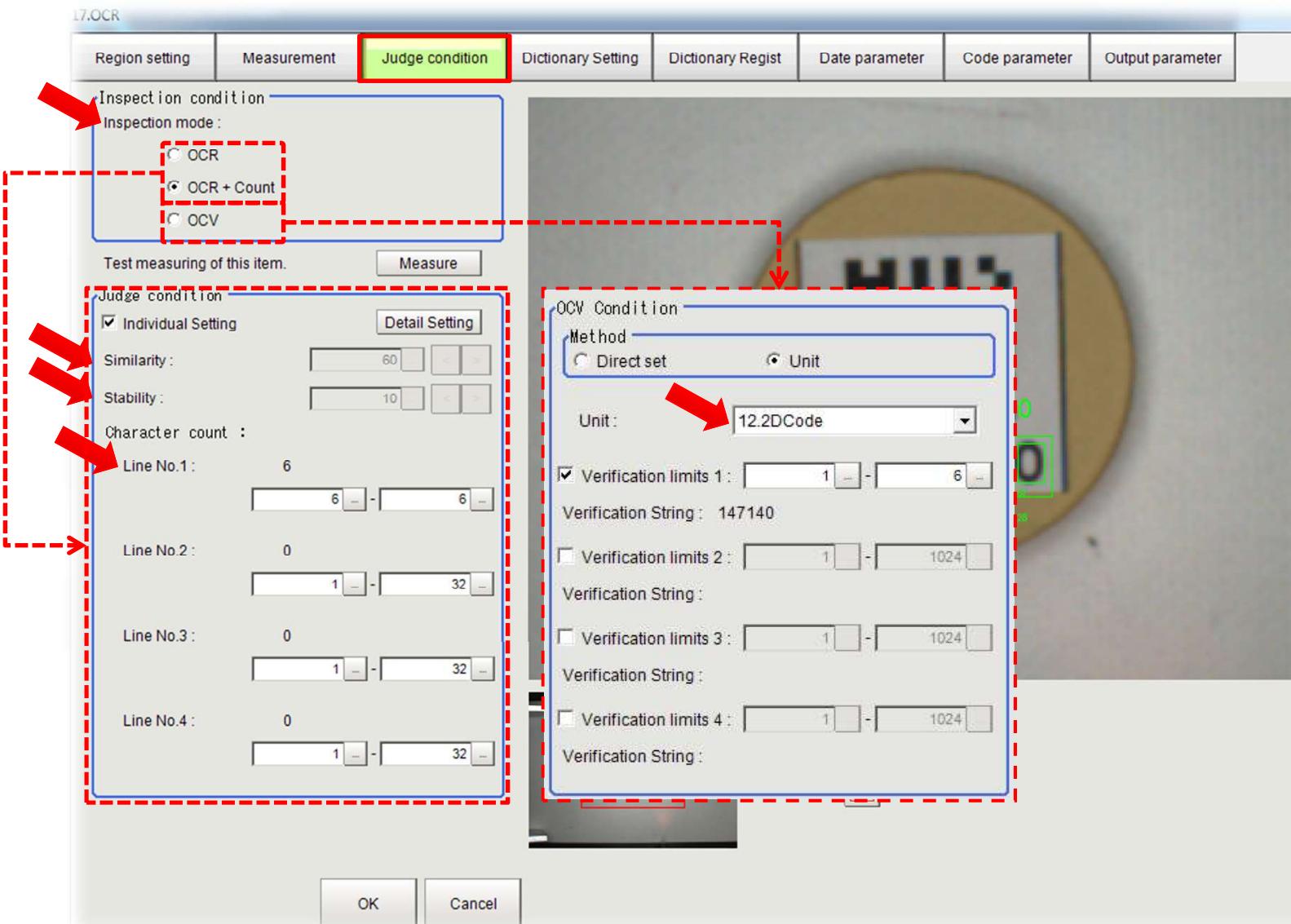
# OCR – Dạy học (Video)

The image is a composite of several logos and text elements:

- Top Left:** Think & See logo with a blue swoosh.
- Top Center:** OCR logo with the text "OCR the OMRON Character Recognition".
- Top Right:** OMRON logo.
- Middle Section:** Sysmac logo with the tagline "always in control".
- Bottom Section:** Think & See logo with a blue swoosh.
- Bottom Left:** Sysmac logo with the tagline "always in control".
- Bottom Center:** Copyright notice: "©OMRON Corporation 2014 All Rights Reserved".

**OMRON**

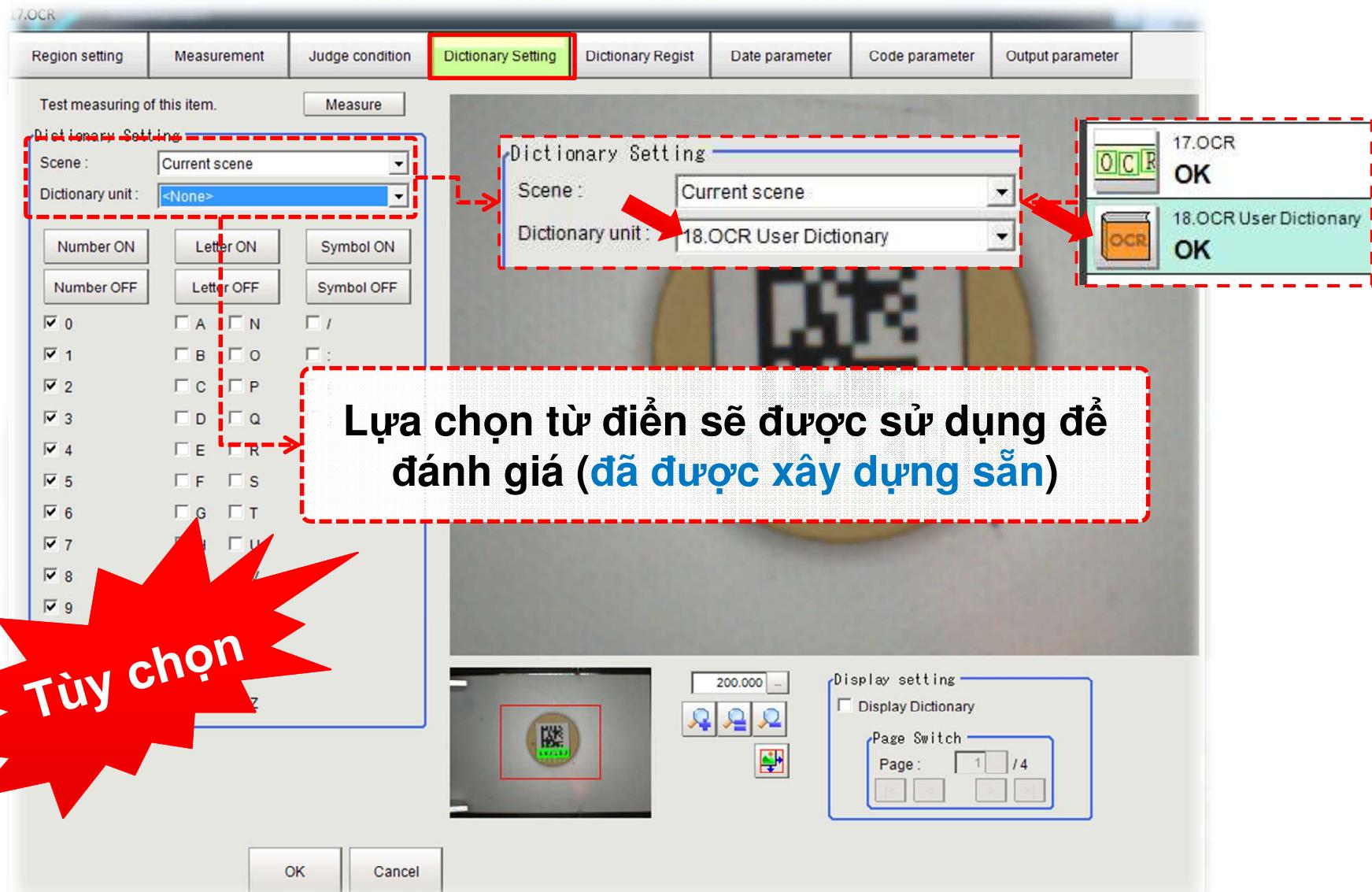
# OCR – Điều kiện đánh giá



OMRON

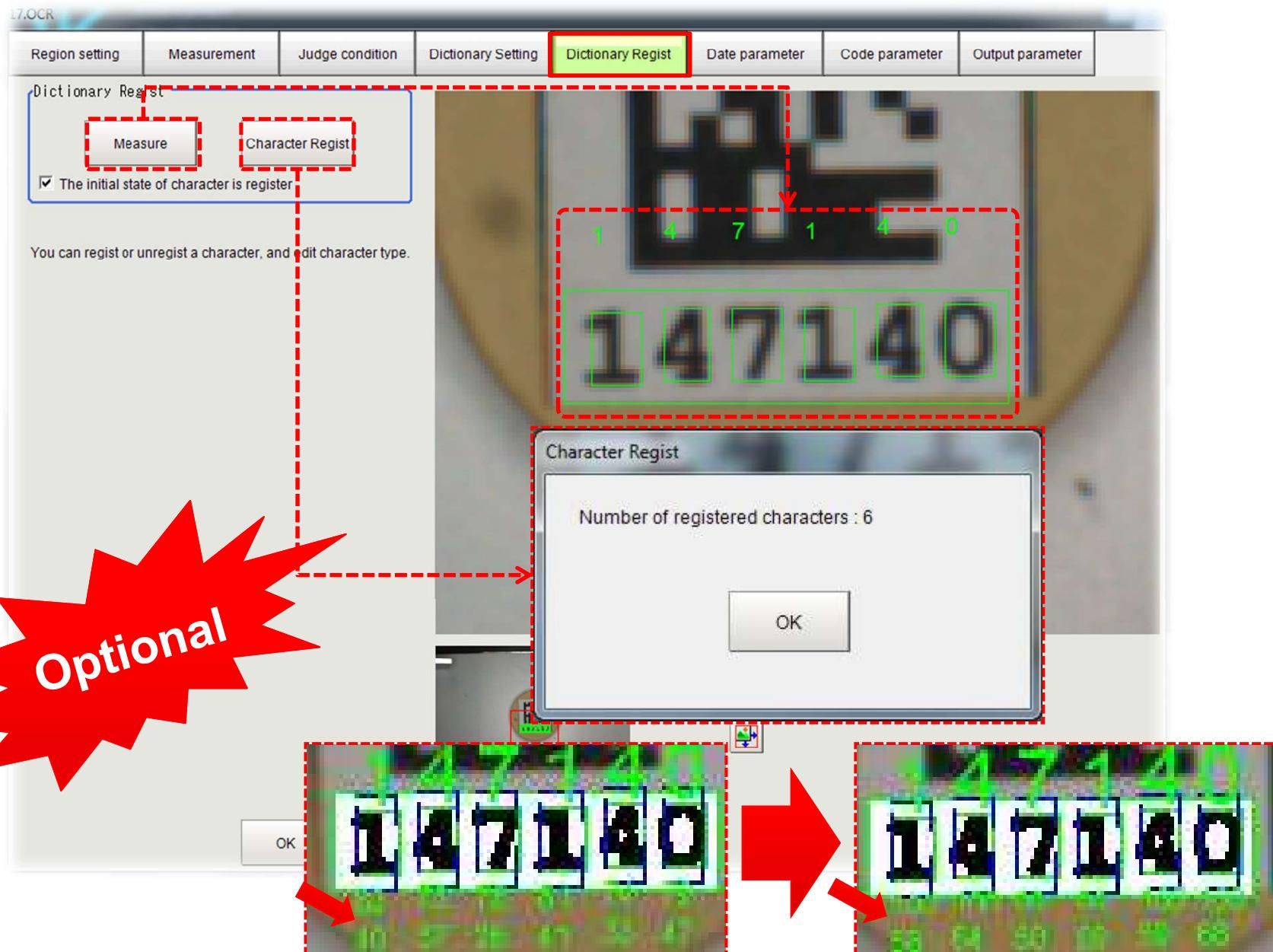


# OCR – Cài đặt

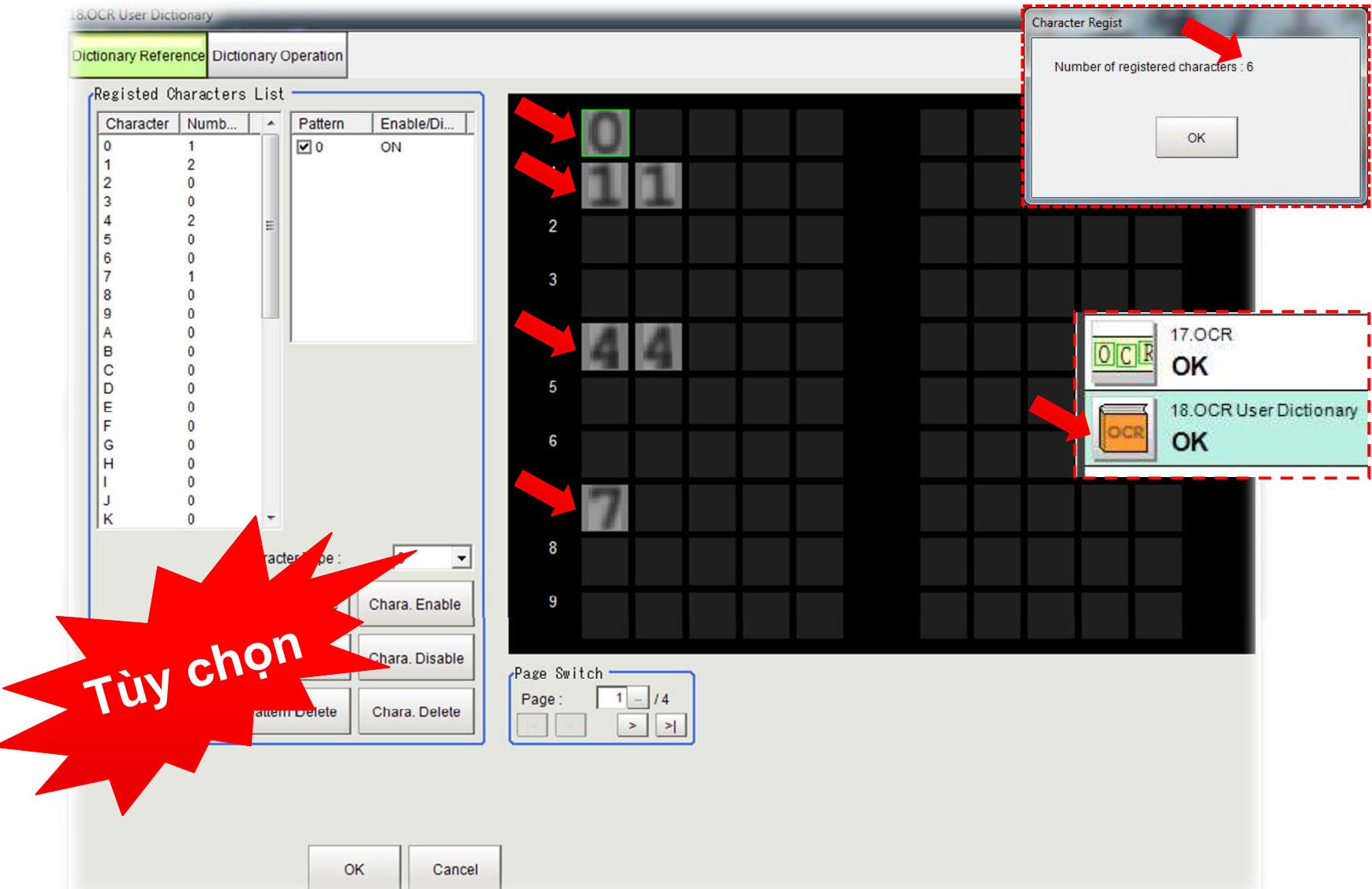


OMRON

# OCR – Đăng ký từ điển



# OCR User Dictionary – Đăng ký mẫu



OMRON

# DEMO: OCR

- Thực hành nhận dạng OCR



# Câu hỏi?

---



**OMRON**