

## TỔNG QUAN PHƯƠNG PHÁP ĐÁNH GIÁ KIỂU HÌNH HIỆU NĂNG CAO TRÊN CÂY TRỒNG: TIẾN TRÌNH PHÁT TRIỂN VÀ TIỀM NĂNG ỨNG DỤNG CHO VIỆT NAM

Nguyễn Trung Đức<sup>1\*</sup>, Phạm Quang Tuấn<sup>1</sup>, Nguyễn Thị Nguyệt Anh<sup>1</sup>, Nguyễn Văn Mười<sup>1</sup>,  
Phùng Danh Huân<sup>1</sup>, Vũ Hải<sup>2</sup>, Trần Văn Quang<sup>3</sup>, Vũ Thị Xuân Bình<sup>4</sup>, Vũ Văn Liết<sup>3</sup>

<sup>1</sup>*Viện Nghiên cứu và Phát triển cây trồng, Học viện Nông nghiệp Việt Nam*

<sup>2</sup>*Viện Nghiên cứu quốc tế về Thông tin đa phương tiện, Truyền thông và Ứng dụng (MICA),  
Đại học Bách Khoa Hà Nội*

<sup>3</sup>*Khoa Nông học, Học viện Nông nghiệp Việt Nam*

<sup>4</sup>*Ban Khoa học và Công nghệ, Học viện Nông nghiệp Việt Nam*

\*Tác giả liên hệ: [ntduc@vnua.edu.vn](mailto:ntduc@vnua.edu.vn)

Ngày nhận bài: 06.08.2021

Ngày chấp nhận đăng: 09.12.2021

### TÓM TẮT

Năng suất cây trồng cần phải tăng gấp đôi hiện tại để đáp ứng nhu cầu của 10 tỉ người tới năm 2050 là một thách thức toàn cầu, đòi hỏi các phương pháp chọn tạo giống mới với hiệu năng và độ chính xác cao. Với sự phát triển của khoa học máy tính, cảm biến hình ảnh, học máy, trí tuệ nhân tạo ngày nay đã giúp các nhà khoa học đánh giá chính xác kiểu hình trong sự tương tác giữa kiểu gen với môi trường ngày càng đa dạng và phức tạp. Đây là nền tảng ra đời kỹ thuật đánh giá kiểu hình thế hệ mới: phương pháp đánh giá kiểu hình cây trồng hiệu năng cao (HTP) kết hợp đa hình ở nhiều cấp độ từ tế bào, cơ quan, cá thể đến cấp độ quần thể cây trồng. Việt Nam là một đất nước thuần nông đang phát triển, chịu ảnh hưởng mạnh của biến đổi khí hậu. Do vậy, việc áp dụng các thành tựu từ phương pháp HTP góp phần rút ngắn thời gian đánh giá, chọn tạo, tạo ra nhiều giống mới thích ứng cao với sự biến đổi khí hậu ngày càng khó lường như hiện nay. Nghiên cứu này trình bày tóm tắt sự ra đời, phát triển, những thách thức đang gặp phải của phương pháp HTP và tiềm năng ứng dụng cho Việt Nam.

Từ khóa: HTP, kiểu hình,  $G \times E \times M$ , dữ liệu lớn, học máy.

### A Review of High-Throughput Crop Phenotyping: Progress and Application for Vietnam

#### ABSTRACT

Double increase in food production to feed 10 billion people sustainably by 2050 is a global challenge, which requires novel breeding methods with high-throughput and accuracy. The development of computer science, image sensors, machine learning, and artificial intelligence provided scientists with new methods for quantitative evaluation of plant phenotypes in the interaction between genotype and environment. It has generated a new area for quantitative analysis of phenotypes: high-throughput crop phenotyping (HTP) combining multidimensional from cellular, tissue, organ, individual to population level. Vietnam is a developing country, agriculture still plays a vital role in economic activity and is strongly influenced by climate change. Therefore, the application of achievements from HTP technology will contribute to shorten the time of evaluation and breeding cycle and develop new resilience varieties highly adaptable to climate change. This study highlighted the history, development and challenges of HTP and its potential application for Vietnam.

Keywords: HTP, phenotyping,  $G \times E \times M$ , big data, machine learning.

#### 1. ĐẶT VẤN ĐỀ

Nhu cầu cung cấp lương thực, thực phẩm, thiếu hụt tài nguyên, biến đổi khí hậu và sử

dụng năng lượng là những thách thức chúng ta phải đối mặt trong sự phụ thuộc vào thực vật. Cho đến năm 2050, sản lượng cây trồng sẽ phải tăng gấp đôi để đáp ứng sự gia tăng dân số toàn

cầu đạt khoảng 10 tỉ người (Hickey & cs., 2019). Cải tiến di truyền vẫn đóng vai trò chính trong việc nâng cao năng suất cây trồng, nhưng tốc độ cải tiến hiện tại không thể đáp ứng nhu cầu về tính bền vững và an ninh lương thực. Chiến lược chọn giống phân tử quan tâm nhiều hơn đến các lựa chọn dựa trên thông tin kiểu gen, tuy nhiên kiểu hình là chìa khóa giải mã vai trò của các gen (Araus & Cairns, 2014).

Sự ra đời của công nghệ giải trình tự gen thế hệ mới (Next-Generation Sequencing), dữ liệu về kiểu gen thu được nhanh chóng với chi phí rẻ hơn (Behjati & Tarpey, 2013). Tuy nhiên, việc sử dụng thông tin trên để thiết lập mối liên hệ toàn bộ gen hiện nay hạn chế do không có sẵn dữ liệu kiểu hình chính xác. Kiểu hình thu thập chỉ ở dữ liệu điểm cuối (năng suất) hoặc dữ liệu kiểu hình một vài điểm (đẻ nhánh, trở bông, chín sữa, chín hoàn toàn) không có tính liên tục. Do đó, không thể nắm bắt được tương tác kiểu gen và môi trường động trên quy mô toàn sinh vật. Ví dụ, cùng một mức năng suất cây trồng có thể đạt được bằng một số kết hợp thay đổi quá trình sinh lý trong các kiểu gen khác nhau. Mặt khác, số lượng tính trạng quan sát và đo đếm được thấp. Những điều này dẫn đến nút thắt trong phân tích tương tác kiểu gen và môi trường.

Thông thường, kích thước quần thể cây trồng của các chương trình chọn giống ứng dụng nhờ chỉ thị phân tử có thể có từ 200 đến tối đa 10.000 dòng, do vậy đánh giá kiểu hình chính xác với hiệu năng cao đồng thời đặc trưng cho hàng trăm dòng thì vẫn còn nhiều thách thức (Araus & Cairns, 2014). Đánh giá kiểu hình theo truyền thống dựa vào con người đòi hỏi nhiều công sức, tốn thời gian và thường có sai số lớn. Kết hợp các công nghệ cảm biến và ứng dụng thuật toán cho việc đánh giá kiểu hình tự động sẽ khắc phục được hạn chế của phương pháp truyền thống. Phương pháp mới này được gọi là phương pháp đánh giá kiểu hình cây trồng hiệu năng cao (HTP) cung cấp các phép đo tự động, có hệ thống, tiết kiệm thời gian cùng với các lợi thế của phép đo, đếm không phá hủy, không xâm lấn, quan sát chính

xác, khách quan, đánh giá thường xuyên theo chu kỳ sinh trưởng và lưu trữ dữ liệu trực tiếp (Zhao & cs., 2019).

Tại Việt Nam, trong lĩnh vực y tế, việc áp dụng các cảm biến như X-quang, nhũ ảnh, cắt lớp vi tính, cộng hưởng từ để thu thập, xây dựng dữ liệu lớn, áp dụng các thuật toán học máy, học sâu, trí tuệ nhân tạo để xử lý và chẩn đoán dựa trên hình ảnh đã có những bước tiến lớn với tiên phong là Viện Nghiên cứu Dữ liệu lớn (VinBigdata) của tập đoàn Vingroup. Chiến lược của VinBigdata là kết hợp thông tin từ hình ảnh và các loại dữ liệu y tế lâm sàng, giải phẫu bệnh, giải mã gen với mục tiêu hướng tới sàng lọc và hỗ trợ chẩn đoán một số bệnh nan y phổ biến ở Việt Nam. Tuy vậy, các nghiên cứu tương tự trong ngành nông nghiệp và ứng dụng trong chọn giống cây trồng để rút ngắn chu kỳ chọn tạo vẫn là một khoảng cách rất lớn. Do đó, nhu cầu cấp thiết cần phải có nghiên cứu tổng quan để tổng hợp, phân tích, đánh giá có hệ thống về phương pháp HTP qua đó thúc đẩy việc ứng dụng HTP trong ngành nông nghiệp nhằm bắt kịp xu hướng nghiên cứu với thế giới.

Mục tiêu của nghiên cứu này cung cấp một cái nhìn tổng quan về HTP bao gồm lịch sử hình thành, nền tảng của phương pháp HTP, ứng dụng HTP trên thế giới và tiềm năng ứng dụng HTP cho Việt Nam. Phương pháp thu thập thông tin cho nghiên cứu tổng quan này dựa trên các kết quả nghiên cứu, các thông tin đã công bố trên tạp chí chuyên ngành uy tín trong danh mục ISI/Scopus và có chỉ số trích dẫn cao. Mười quy tắc để soạn một bài tổng quan đề xuất bởi Pautasso (2013) đã được áp dụng trong nghiên cứu này. Phương pháp tổng hợp, kế thừa và phương pháp nghiên cứu tại bàn đã được áp dụng để tổng hợp, phân tích lịch sử ra đời, ứng dụng của phương pháp HTP trên thế giới qua đó đánh giá, nhận định và đề xuất chiến lược ứng dụng cho Việt Nam.

## 2. LỊCH SỬ HÌNH THÀNH

Lịch sử quá trình hình thành và phát triển của các phương pháp HTP có thể được chia

thành ba giai đoạn chính (Bảng 1). Giai đoạn sơ khai từ năm 1911 đến năm 1997 là thời kỳ hình thành khái niệm kiểu hình và kiểu hình học - “phenomics”. Thuật ngữ “kiểu hình” như một khái niệm song hành với “kiểu gen” được tạo ra cách đây một thế kỷ (Johannsen, 1911), được sử dụng để mô tả một loạt các đặc điểm ở thực vật, vi khuẩn, nấm và động vật. Kiểu hình thực vật là chức năng cơ thể thực vật được hình thành trong quá trình sinh trưởng và phát triển của thực vật từ mối quan hệ tương tác động giữa nền gen (kiểu gen) và thể giới vật chất mà thực

vật phát triển (môi trường). Thuật ngữ “kiểu hình” được bắt đầu được sử dụng vào những năm 1960 (Walter & cs., 2015) và sau đó được gọi là tập hợp các phương pháp và quy trình được sử dụng để đo lường chính xác sự phát triển, cấu trúc và thành phần của cây trồng ở các quy mô khác nhau (Fiorani & Schurr, 2013). Đến năm 1997 thì khái niệm kiểu hình học - “phenomics” được đề xuất bởi Schork (1997) khi tác giả tìm hiểu cách tiếp cận, vấn đề và giải pháp trong nghiên cứu di truyền các bệnh phức tạp trên người.

**Bảng 1. Lịch sử của phương pháp HTP trên cây trồng**

Giai đoạn	Năm	Nội dung	Nguồn trích dẫn
Giai đoạn sơ khai: thời kỳ hình thành khái niệm của kiểu hình và HTP	1911	Khái niệm kiểu hình lần đầu tiên được đề xuất bởi nhà di truyền học người Đan Mạch Johannsen	Johannsen (1911)
	1997	Khái niệm về kiểu hình học - “phenomics” được đề xuất	Schork (1997)
Giai đoạn phát triển mạnh: từ cuối thế kỷ XX, các nhóm nghiên cứu kiểu hình thực vật và các tổ chức thương mại đã được thành lập liên tiếp và một loạt các công cụ kiểu hình thông lượng cao, độ chính xác cao, tự động hoặc bán tự động đã được phát triển để có được kiểu hình thực vật chất lượng cao, có thể lập lại dữ liệu	1998	Phát triển nền tảng HTP trên quy mô lớn	<a href="http://www.cropdesign.com">http://www.cropdesign.com</a>
	2007	Trung tâm nghiên cứu HTP đầu tiên tại Úc được thành lập và đặt tên là The Australian Plant Phenomics Facility (APPF)	<a href="https://www.plantphenomics.org.au">https://www.plantphenomics.org.au</a>
	2011	Thảo luận về nút thắt giữa kiểu hình và kiểu gen cùng các vấn đề cần được giải quyết	Furbank & Tester (2011)
	2012	Mạng kiểu hình thực vật châu Âu (EPPN) được thành lập, hoàn thành dự án nghiên cứu chung EPPN giai đoạn 2012-2015 và tiếp tục với các chương trình EPPN2020 và EMPHASIS với mục tiêu phát triển cơ sở hạ tầng HTP dài hạn toàn châu Âu	<a href="https://eppn2020.plant-phenotyping.eu">https://eppn2020.plant-phenotyping.eu</a> <a href="https://emphasis.plant-phenotyping.eu">https://emphasis.plant-phenotyping.eu</a> ;
		Nhiều thuật toán và công cụ mới đã được đề xuất để xử lý các đặc điểm vi mô của rễ, thân và hạt, như RootAnalyzer, VesselParser	Burton & cs. (2012)
		Đề xuất rằng trong tất cả các nghiên cứu di truyền các tính trạng số lượng (QTLs) thì kiểu hình là vua và kiểu gen là nữ hoàng	Tuberosa (2012)
Giai đoạn phát triển có hệ thống: đang bước vào kỷ nguyên mới tên là phenomics, cung cấp dữ liệu lớn và hỗ trợ quyết định cho việc tiết lộ cơ chế phân tử và chức năng gen của thực vật	2013	Kiểu hình nên được liên kết chặt chẽ với các công nghệ, như lập bản đồ liên kết gen độ phân giải cao, nghiên cứu liên kết toàn hệ gen và mô hình chọn lọc dựa trên bộ gen	Cobb & cs. (2013)
	2016	Phát triển nền tảng HTP đầu tiên trên đồng ruộng, mở ra kỷ nguyên từ cảm biến tới cây trồng	<a href="http://www.lemnatec.com">http://www.lemnatec.com</a>
		Đăng ký mạng lưới kiểu hình thực vật quốc tế (IPPN)	<a href="https://www.plantphenotyping.org">https://www.plantphenotyping.org</a>
	2017	Chiến lược cho HTP đa quy mô, đa miền được đề xuất	Tardieu & cs. (2017)
	2019	Một số mạng lưới kiểu hình thực vật quốc gia và khu vực (PPNs) đã được tổ chức, chẳng hạn như FPPN, PPA, NAPPN, CPPN,...	Carroll & cs. (2019)
2021	HTP đóng vai trò không thể thiếu trong phát triển nhanh chóng các giống cây trồng đảm bảo an ninh lương thực toàn cầu	Varshney & cs. (2021)	

Giai đoạn hai là thời kì phát triển mạnh của phương pháp HTP từ năm 1998 đến năm 2012. Từ cuối thế kỷ XX, các nhóm nghiên cứu kiểu hình thực vật và các tổ chức thương mại đã được thành lập liên tiếp và một loạt công cụ kiểu hình thông lượng cao, độ chính xác cao, tự động hoặc bán tự động đã được phát triển để có được kiểu hình thực vật chất lượng cao, có thể lặp lại dữ liệu. Thực tế cho thấy, so với các dữ liệu lớn về bộ gen đã được giải trình tự thì vẫn có một nút thắt về thông tin kiểu hình cản trở tiến trình tìm hiểu cơ sở di truyền của các tính trạng phức tạp (Furber & Tester, 2011). Để phá vỡ nút thắt này và cải thiện hiệu quả chọn lọc ở cấp độ phân tử, các công nghệ HTP sẽ cung cấp cho nhà chọn giống những hiểu biết mới trong việc lựa chọn các loài mới để thích ứng với sự suy giảm nguồn tài nguyên thiên nhiên và biến đổi khí hậu trên toàn thế giới. Trong kỷ nguyên ứng dụng và phân tích bộ gen cây trồng, Tuberosa (2012) đã nhấn mạnh tầm quan trọng như nhau của cả kiểu gen và kiểu hình cũng như khả năng chống chịu điều kiện bất thuận của cây trồng.

Giai đoạn ba từ năm 2013 đến nay là giai đoạn phát triển HTP có hệ thống từ nhà kính đến thực địa ngoài đồng ruộng. Những tiến bộ đồng thời trong việc đánh giá kiểu hình thực vật chủ yếu nhờ vào những đổi mới trong công nghệ hình ảnh và cảm biến đã được kết hợp với dữ liệu bộ gen để lập bản đồ liên kết gen độ phân giải cao, nghiên cứu liên kết toàn hệ gen và mô hình chọn lọc dựa trên bộ gen (Cobb & cs., 2013). Phát triển nền tảng HTP đầu tiên trên đồng ruộng, mở ra kỷ nguyên từ cảm biến tới cây trồng. Mạng lưới kiểu hình thực vật quốc tế (IPPN) đã được đăng ký. Thế giới đang bước vào kỷ nguyên mới có tên là phenomics - hệ thống đánh giá HTP, cung cấp dữ liệu lớn và hỗ trợ quyết định cho việc tiết lộ cơ chế phân tử và chức năng gen của thực vật (Tardieu & cs., 2017; Carroll & cs., 2019). Khoảng 20% các nền tảng hệ thống HTP đã được thành lập trên toàn thế giới (Yang & cs., 2020). Gần đây, Varshney & cs. (2021) đã khẳng định HTP đóng vai trò không thể thiếu trong chiến lược phát triển nhanh các giống cây trồng đảm bảo an ninh lương thực toàn cầu.

### 3. NỀN TẢNG CỦA PHƯƠNG PHÁP ĐÁNH GIÁ KIỂU HÌNH HIỆU NĂNG CAO

Sản xuất cây trồng rất phức tạp và được xác định bởi nhiều yếu tố bao gồm kiểu gen, môi trường trồng trọt (ví dụ: thời tiết, đất, vi khí hậu và vị trí) và các phương pháp canh tác (phân bón và quản lý dịch hại). Tất cả các tác động của các yếu tố khác nhau đến sản xuất cây trồng có thể được tóm tắt bằng cách sử dụng một phương trình sản xuất cây trồng, tức là, năng suất cây trồng (P) là hàm số của các tương tác giữa kiểu gen (G), môi trường (E) và phương pháp canh tác (M) được minh họa ở công thức dưới đây:

$$P = G \times E \times M + \varepsilon$$

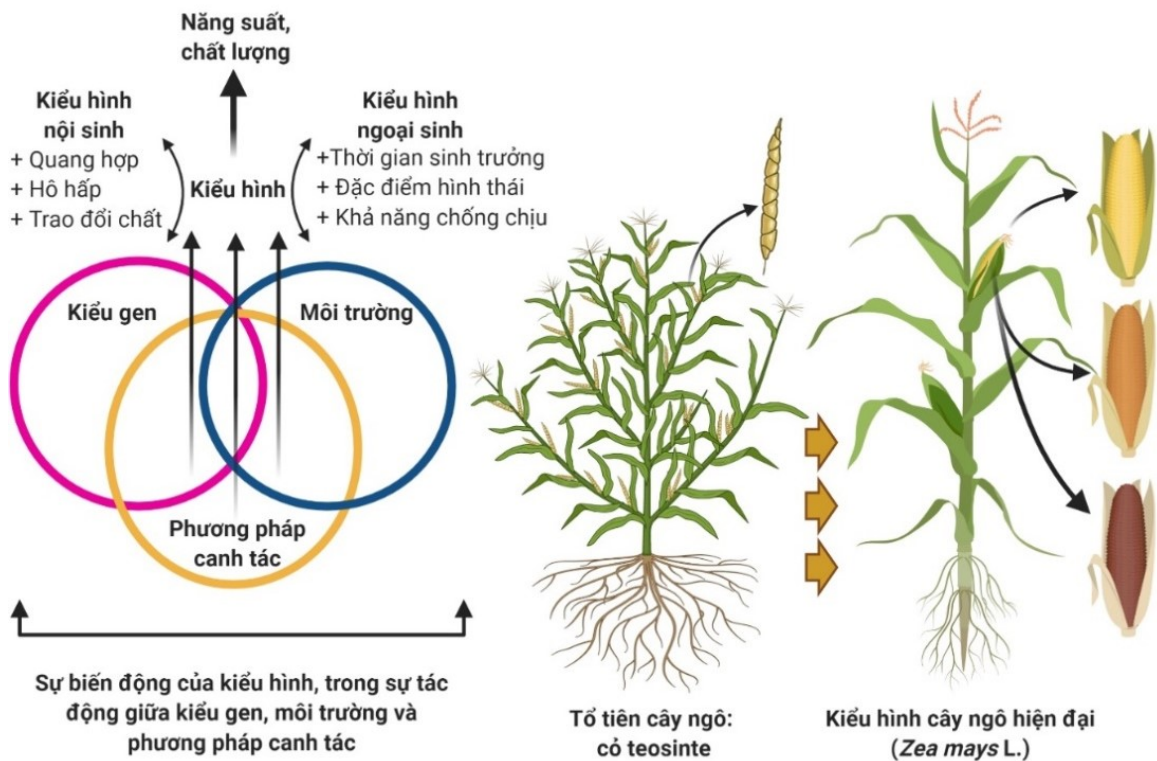
Trong đó, P là kiểu hình thực vật đề cập đến các đặc tính vật lý có thể quan sát được của một sinh vật, bao gồm năng suất; G là kiểu gen liên quan đến cấu tạo di truyền của một sinh vật; E là các yếu tố môi trường ảnh hưởng đến sự phát triển của thực vật bao gồm khí hậu, chất lượng đất, ánh sáng, nhiệt độ và lượng nước sẵn có; M là phương pháp canh tác, thực hành quản lý cây trồng và đồng ruộng, chẳng hạn như xử lý hạt giống, trồng trọt, quản lý dịch hại, quản lý dinh dưỡng và tưới tiêu;  $\varepsilon$  là tổng sai số của mô hình. Công thức này được đề xuất bởi Beres & cs. (2020) trong nghiên cứu đề xuất chiến lược và sáng kiến lúa mì toàn cầu. Phương trình cho thấy kiểu hình cây trồng chịu tác động bởi kiểu gen và các yếu tố ngoại cảnh, phương pháp canh tác (Hình 1). Môi trường thay đổi, cây trồng cũng không ngừng tiến hóa và thay đổi theo. Ví dụ, từ cây cỏ teosinte nhiều nhánh qua thời gian dài dưới ảnh hưởng của môi trường và chọn lọc từ con người đã tiến hóa thành cây ngô hiện đại (*Zea mays* L.) như ngày nay (Hình 1).

Các phương pháp đánh giá kiểu hình cây trồng truyền thống tốn nhiều công sức, thời gian, chủ quan và thường xuyên phải thu mẫu cây trồng trực tiếp (Furber & Tester, 2011). Sự chậm trễ trong việc ứng dụng tiến bộ của các công nghệ mới nổi và năng suất thấp trong việc đánh giá kiểu hình cây trồng đã trở thành hạn chế lớn đối với nghiên cứu về gen chức

năng và chọn giống cây trồng (Deery & cs., 2016). Công nghệ HTP nổi lên trong thập kỷ qua nhờ những tiến bộ và giảm chi phí trong công nghệ cảm biến, thị giác máy tính, tự động hóa và các mô hình học máy tiên tiến. HTP đề cập đến việc thu thập dữ liệu kiểu hình đa chiều ở nhiều cấp độ từ tế bào, cơ quan, thực vật đến quần thể bằng cách sử dụng các công nghệ mới (Zhao & cs., 2019). Một hệ thống HTP toàn diện bao gồm phần cứng (cảm biến và nền tảng) và phần mềm tính toán (lưu trữ, trích xuất và phân tích dữ liệu). Các cảm biến được sử dụng rộng rãi trong công nghệ HTP chủ yếu là các cảm biến không tiếp xúc và không xâm lấn, chẳng hạn như máy ảnh kỹ thuật số RGB, đa phổ, siêu phổ, cảm biến nhiệt và cảm biến ba chiều LiDAR (Araus & cs., 2018; Zhao & cs., 2019). Một số nghiên cứu hiện nay đã thử nghiệm và áp dụng một số kỹ thuật hình ảnh tiên tiến được sử dụng rộng rãi

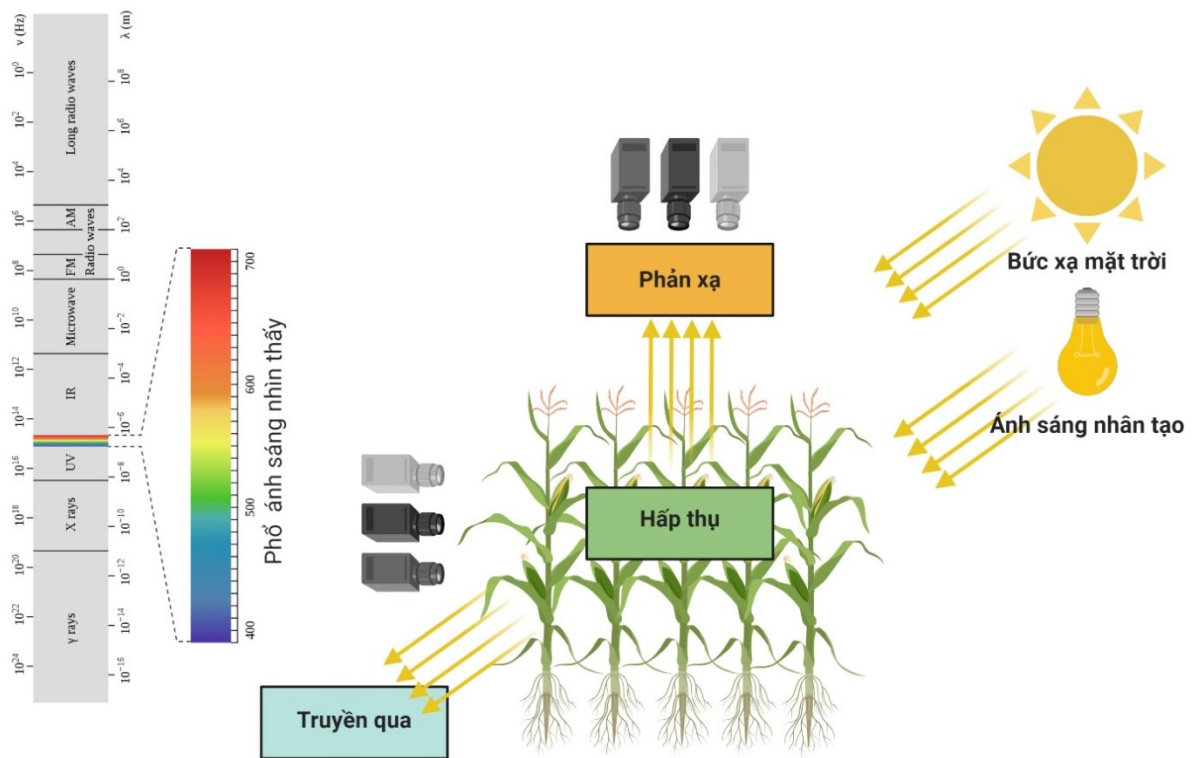
trong các ứng dụng y tế, chẳng hạn như chụp cộng hưởng từ (MRI), chụp cắt lớp phát xạ positron (PET) và chụp cắt lớp vi tính (CT) cho các hệ thống HTP đang phát triển bùng nổ mô nhỏ hoặc nhà kính (Yang & cs., 2020). Những tiến bộ trong công nghệ cảm biến chủ yếu được thúc đẩy bởi lĩnh vực công nghiệp. Nhìn thấy tiềm năng ứng dụng của chúng, rất nhiều các nhà khoa học trên thế giới đã và đang nghiên cứu để tích hợp chúng vào hệ thống HTP ứng dụng trên cây trồng.

Dựa vào sự hấp thụ (absorbance), phản xạ (reflectance) và truyền qua vật (transmittance) của ánh sáng ở bước sóng từ đến tia gamma ( $10^{-12}m$ ) đến sóng radio ( $10^6m$ ), vô số cảm biến được phát triển như VIS (400-750nm), VNIR (400-1.000nm), SWIR (1.000-2.500nm) (Hình 2). Ứng dụng phần mềm, công cụ trích xuất hình ảnh (Bảng 2) mà vô số tính trạng kiểu hình được trích xuất.



Ghi chú: Hình được thiết kế trên ứng dụng trực tuyến <https://app.biorender.com/>, các icon được sử dụng từ các nguồn có sẵn trên ứng dụng.

**Hình 1. Sự biến động của kiểu hình trong sự tương tác giữa kiểu gen, môi trường và phương pháp canh tác**



Ghi chú: Hình được thiết kế trên ứng dụng trực tuyến <https://app.biorender.com/>, các icon được sử dụng từ các nguồn có sẵn trên ứng dụng.

**Hình 2. Nguyên lý của phương pháp đánh giá kiểu hình không phá hủy**

**Bảng 2. Công cụ trích xuất và xử lý hình ảnh**

Công cụ	Mô tả	Nguồn trích dẫn
SmartRoot	Phân tích cấu trúc bộ rễ	Lobet & cs. (2011)
RhizoVision Analyzer	Phân tích cấu trúc bộ rễ	Seethepalli & cs. (2020)
HTPheno	Shoot and plant architecture	Hartmann & cs. (2011)
Canopy Reconstruction	Công cụ tái cấu trúc tán cây ở lúa mì và lúa nước	Pound & cs. (2014)
Integrated Analysis Platform	Mã nguồn mở xây dựng trên nền tảng JAVA với gần 33,700 ảnh gốc phục vụ nghiên cứu.	Klukas & cs. (2014)
PlantCV	Xử lý hình ảnh, trích xuất số liệu chạy trên nền tảng Python miễn phí	Gehan & cs. (2017), <a href="https://plantcv.danforthcenter.org">https://plantcv.danforthcenter.org</a>
Image Harvest	Xử lý hình ảnh từ hệ thống đánh giá kiểu hình hiệu năng cao chạy trên nền tảng Python miễn phí	Knecht & cs. (2016)
LemnaGrid	Công cụ trả phí kèm theo hệ thống Scanalyzer (Lemna Tec) cực kì mạnh mẽ.	<a href="https://www.lemnatec.com/">https://www.lemnatec.com/</a>
SmartGrain	Phân tích hình dạng và kích thước hạt lúa (đậu tương, dưa ...)	Tanabata & cs. (2012)
GrainScan	Phân tích hình dạng và kích thước hạt lúa mì	Whan & cs. (2014)

Ghi chú: Cập nhật công cụ phân tích và nguồn trích dẫn được trên nền tảng web <https://www.quantitative-plant.org>.

Ví dụ minh họa cho các bước thu thập và xử lý hình ảnh cây trồng quy mô nhà kính được mô

tả ở hình 3. Ảnh gốc sau khi chụp bằng RGB camera được xử lý có hệ thống bằng phần mềm

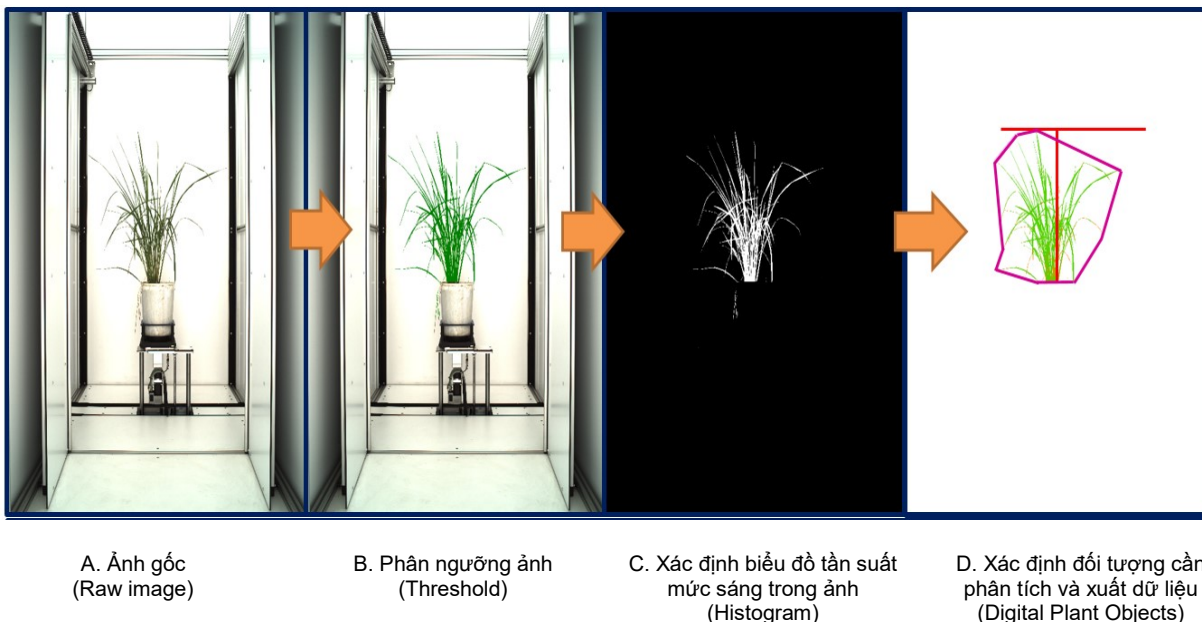


LemnaGrid (Lemna Tec, Germany). Đầu tiên, ảnh gốc được phân ngưỡng ảnh, sau đó áp dụng biểu đồ tần suất để thống kê số lần xuất hiện các mức sáng trong ảnh, từ các thông số này xác định đối tượng cần phân tích, các thông tin cần trích xuất và xuất file dữ liệu sang dạng “.csv” (Hình 3). Các tính trạng trích xuất bao gồm chiều cao cây, số lá, góc lá, diện tích lá, cấu trúc tán cây, khối lượng tươi, khối lượng chất khô, các chỉ số sinh lý dưới dạng số với đơn vị là pixel. Mô hình tiên lượng được xây dựng trên một tập con khoảng 200-300 cá thể cây trồng thông qua mô hình tuyến tính đơn biến hoặc qua các mô hình học máy phức tạp hơn để tìm hiểu sự tương quan giữa các tính trạng số và giá trị đo đếm thực tế. Các tính trạng số trên một quần thể lớn thông qua các mô hình tiên lượng sẽ được chuyển đổi sang giá trị kiểu hình thực tế với độ chính xác rất cao.

#### 4. ỨNG DỤNG HTP TRÊN THẾ GIỚI

Các nền tảng đánh giá kiểu hình hiệu năng cao quy mô nhà kính và ngoài đồng ruộng, trong cả môi trường học thuật và công nghiệp, đã phát triển nhanh chóng trong 10 năm qua và đang

tiếp tục được cải thiện. Ví dụ, một cơ sở đánh giá kiểu hình hiệu năng cao dựa trên nhà kính đã được phát triển ở Trung Quốc, bao gồm một người ghi điểm tính năng suất mới có thể tự động đo tổng số hạt giống, số hạt đầy, khả năng sinh sản của hạt, năng suất trên mỗi cây, trọng lượng hạt, hình dạng và kích thước hạt với tốc độ một cây mỗi phút (Yang & cs., 2014). Gần đây, các cảm biến siêu phổ và phản xạ đa năng gắn trên máy kéo đã được sử dụng để đánh giá kiểu hình trên một quần thể gồm 1.516 dòng lúa tái tổ hợp (Tanger & cs., 2017). So với việc đo đạc bằng tay, hệ thống máy kéo đã thu thập được kiểu hình có hiệu năng và độ chính xác cao như chiều cao cây, thời gian ra hoa, năng suất hạt và chỉ số thu hoạch. Các hệ thống hiệu năng cao khác bao gồm các nền tảng giàn, chẳng hạn như Máy quét trường TERRA-REF ở Arizona (Danilevicz & cs., 2021), hệ thống kiểu hình LeasyScan (Vadez & cs., 2015) và máy bay không người lái (UAV) (Shi & cs., 2016) tại Viện Nghiên cứu Cây trồng quốc tế cho vùng nhiệt đới bán khô cằn (ICRISAT) ở Ấn Độ. Yang & cs. (2020) đã trình bày rất chi tiết ưu điểm và hạn chế của các nền tảng ứng dụng HTP được trình bày tại bảng 3.



Nguồn: Nguyen (2020).

Hình 3. Các bước chính phân tích hình ảnh trên cây lúa

**Bảng 3. Ưu điểm và hạn chế của các nền tảng ứng dụng HTP**

Ứng dụng	Nền tảng	Ưu điểm	Hạn chế
Đánh giá cấu trúc thân cây Arabidopsis trong phòng thí nghiệm	PHENOPSIS (DB); GROWSCREEN (FLUORO); Phenoscope; Phenovator; PlantScreen™	Chi phí phải chăng, nhanh chóng và tự động cho các quần thể có cỡ mẫu lớn	Không phù hợp cho các cây trồng có kích thước lớn
Đánh giá cấu trúc thân cây trồng trong phòng thí nghiệm	TraitMill™; Scanalyzer3D™; PHENOARCH; HRPF	Quan sát được sự phát triển chồi, sinh khối và nhiều thông tin cho các quần thể có cỡ mẫu lớn	Chi phí lớn; cần có các chuyên gia đa ngành duy trì và cập nhật
	PhenoBox	Chi phí phải chăng, dễ vận hành	Cần nhiều lao động để sàng lọc quy mô lớn
Đánh giá cấu trúc rễ trong phòng thí nghiệm	PlaRoM; Rhizoslides; Rhizoponics; RADIX; RhizoTubes	Chi phí phải chăng, thu được mô hình cấu trúc rễ dạng hai chiều	Bộ rễ phải được sinh trưởng ở môi trường trong suốt
	GiARoots RootReader3D	Chi phí phải chăng, thu được mô hình cấu trúc rễ dạng ba chiều	Bộ rễ phải được sinh trưởng ở môi trường trong suốt
	GROWSCREEN-Rhizo	Đánh giá các đặc điểm của cả thân và rễ	Bộ rễ bị hạn chế ở một chiếc hộp có kích thước cố định
	MRI-PET, PET-CT, MRI-CT	Thu được hệ thống rễ 3 chiều	Chi phí cao; tốn nhiều thời gian
Đánh giá kiểu hình dựa trên mặt đất với quy mô đồng ruộng	CPRS - một tháp đánh giá kiểu hình cố định	Dễ cài đặt và bảo trì	Thông tin cây trồng hạn chế ở các khu vực cố định đã được thu thập
	Field Scanalyzer - hệ thống HTP chạy trên đường ray.	Tích hợp các cảm biến quang học khác nhau; độ phân giải hình ảnh cao	Chi phí cao; vùng hình ảnh hạn chế; phụ thuộc vào ánh sáng tự nhiên
	BreedVision - máy kéo tự hành được trang bị nhiều cảm biến gắn trong một buồng tối di động	Tích hợp các cảm biến quang học khác nhau; tình trạng hình ảnh ổn định; không hạn chế khu vực hình ảnh	Hạn chế bởi điều kiện thời tiết (độ ẩm đất, nước, gió, mưa)
Viễn thám với quy mô đồng ruộng	Máy bay không người lái được trang bị nhiều cảm biến	Không hạn chế khu vực chụp ảnh; các phép đo nhanh chóng (sinh trưởng cây trồng, năng suất, phản ứng với stress...); dễ cài đặt và sử dụng	Không thể lấy thông tin bên dưới tán cây; Điều hành nghiêm ngặt và phải tuân thủ luật bay của địa phương
Đánh giá kiểu hình nhỏ gọn	PocketPlant3D - điện thoại thông minh được trang bị ứng dụng để đo cấu trúc tán và đặc điểm lá ngô	Chi phí phải chăng; sử dụng linh hoạt; có thể được phổ biến rộng rãi	Các tính trạng bị hạn chế; thiếu các mô hình mạnh mẽ thay đổi sao cho phù hợp với các điều kiện phức tạp tại thực địa
Đánh giá kiểu hình sau thu hoạch	Máy gia tốc đánh giá hạt giống (SEA)	Tự động tuốt bông lúa và đo nhanh các tính trạng cấu thành năng suất	Không thể có được các đặc điểm hạt ở dạng 3 chiều và đặc điểm bông
	P-TRAP; PANorama	Định lượng cả tính trạng bông lúa và tính trạng hạt; không cần tách hạt	Cần tách các bông lúa bằng tay; không đo được các đặc điểm hạt dạng ba chiều
	PhenoSeeder	Trích xuất các đặc điểm ba chiều từ các hạt riêng lẻ với độ chính xác cao	Tốc độ đo thấp; cần phải tách hạt
	X-ray CT	Trích xuất các đặc điểm hạt ngũ cốc ba chiều và các đặc điểm bông; không cần tách hạt	Chi phí cao; mất thời gian; cần phát triển chương trình phân tích hình ảnh riêng cho các loài mới
	Hyperspectral imaging	Ước lượng hàm lượng protein và thông tin sinh lý hoặc sinh hóa khác	Chi phí cao; cần phân tích hình ảnh riêng và mô hình cập nhật cho các loài mới, các chỉ số sinh lý hoặc sinh hóa mới

Ứng dụng của các phương pháp đánh giá kiểu hình hiệu năng cao đã được Zhao & cs. (2019) và Yang & cs. (2020) tổng hợp, phân tích

khá đầy đủ. Trong bài tổng quan này, nhóm tác giả chia thành bốn nhóm ứng dụng chính của HTP như sau:



Nhóm 1: HTP cung cấp phép đo hiệu quả và khách quan về các tính trạng cây trồng. Hệ thống HTP có thể đánh giá kiểu hình trên cánh đồng nhân giống một cách hệ thống, hiệu quả hơn và tiết kiệm chi phí hơn, điều này cho phép tăng hiệu suất của các chương trình chọn giống để sàng lọc một quần thể cây trồng có cỡ mẫu lớn và cải thiện cường độ chọn lọc. Ví dụ, nền tảng HTP dựa trên UAV có thể sàng lọc các cánh đồng chọn giống trong một khoảng thời gian ngắn. Thiết bị phản xạ quang phổ và thị giác máy tính cung cấp các tiêu chí nhất quán để ước tính các tính trạng cây trồng theo nhiều chiều, chẳng hạn như chiều cao cây, nhiệt độ cây và hàm lượng diệp lục.

Nhóm 2: HTP giúp xác định các tính trạng cây trồng mới. Các cảm biến tiên tiến (siêu phổ và hồng ngoại) ghi lại thông tin bước sóng phản xạ ngoài tầm nhìn và cảm nhận của con người. Phương pháp phân tích dữ liệu tiên tiến và mô hình trí tuệ nhân tạo (AI) sẽ giúp phát hiện những thông tin ẩn từ dữ liệu đồng thời có tiềm năng lớn trong việc khám phá đặc điểm cây trồng mới. Đặc điểm mới có thể dùng để mô tả năng suất cây trồng ở một giai đoạn sinh trưởng cụ thể hoặc để xác định phản ứng động của cây trồng đối với môi trường theo chu kỳ sinh trưởng. Đặc điểm cây trồng mới phát sinh từ các biến dị có thể cung cấp thêm thông tin để định lượng các biến thể di truyền đặc biệt và có khả năng làm tăng phương sai di truyền.

Nhóm 3: HTP tích hợp dữ liệu kiểu hình và dữ liệu kiểu gen. Các kiểu hình dựa trên HTP có thể được tích hợp vào phân tích di truyền, chẳng hạn như lập bản đồ locus tính trạng số lượng (QTL) hoặc nghiên cứu liên kết toàn hệ gen (GWAS) để xác định các tính trạng di truyền quan trọng. Các gen, QTL triển vọng có thể được áp dụng thông qua chọn lọc có hỗ trợ của chỉ thị phân tử (MAS) để sàng lọc và phát triển các nguồn vật liệu cây trồng mới trong quá trình lai tạo. Việc tích hợp sẽ cho phép chọn lọc chính xác, giảm chu kỳ nhân giống và tăng lợi ích di truyền.

Nhóm 4: HTP cho phép các mô hình hóa  $G \times E \times M$  để dự đoán P chính xác hơn. Công nghệ HTP cho phép thu thập dữ liệu lớn về cây trồng ở độ phân giải cao, đa chiều và khám phá các đặc điểm cây trồng mới để làm sáng tỏ các tương tác

$G \times E \times M$ . Các mô hình, thuật toán tiên tiến dựa trên công nghệ học máy và học sâu, AI sẽ chuyển đổi chương trình chọn giống từ mô tả kiểu hình sang dự đoán kiểu hình cho phép thiết kế các đặc điểm cây trồng dựa trên nhu cầu.

## 5. ỨNG DỤNG HTP CHO VIỆT NAM

Đánh giá kiểu hình vẫn là nền tảng của quá trình chọn tạo giống cây trồng. Mặc dù có những tiến bộ trong di truyền học và ứng dụng công nghệ phân tử, việc chọn giống vẫn dựa chủ yếu vào sự biểu hiện của năng suất và một số đặc điểm nông học quan trọng để chọn và xác định sản phẩm thương mại. Trong điều kiện Việt Nam hiện nay, chúng ta có thể áp dụng ngay một số thành quả nghiên cứu HTP, các phương pháp đánh giá kiểu hình hiệu năng cao, đánh giá cấu trúc bộ rễ, định lượng một số tính trạng như khối lượng chất khô, chiều cao cây, diện tích lá qua ảnh chụp RGB, định lượng số lượng, kích thước hạt, cấu trúc bông, bắp, dự đoán diện tích, năng suất qua ảnh chụp từ flycam với chi phí thấp.

### 5.1. Đánh giá cấu trúc bộ rễ

Xác định đặc điểm cấu trúc bộ rễ đóng vai trò quan trọng trong tính chống chịu đối với các yếu tố bất thuận của môi trường. Cấu trúc bộ rễ cần thiết cho việc thu nhận nước và dinh dưỡng, tương tác vi khuẩn, lưu trữ chất dinh dưỡng, là nơi neo đậu của cây trồng, chống đổ và tác động đến năng suất hạt (Lynch, 2019). WinRHIZO được biến đổi là một phần mềm thương mại dùng để đánh giá cấu trúc bộ rễ qua hình ảnh hiệu quả. Tuy nhiên, hạn chế của phần mềm này là hiệu năng thấp và người dùng phải trả phí bản quyền hàng năm. Falk & cs. (2020a) đã phát triển phần mềm ARIA2.0 (phân tích hình ảnh rễ tự động hiệu năng cao) thay thế phần mềm thương mại WinRHIZO. Hệ thống ARIA2.0 cung cấp một phương pháp đánh giá kiểu hình bộ rễ hiệu năng cao, hiệu quả về chi phí, không phá hủy, cung cấp dữ liệu chuỗi thời gian có liên quan về mặt sinh học về sự tăng trưởng và phát triển của rễ cho các ứng dụng trong chọn giống cây trồng. Falk & cs. (2020b) đã ứng dụng thành công ARIA2.0

trong chọn lọc đánh giá đa dạng di truyền của 292 dòng đậu tương.

### **5.2. Định lượng năng suất sinh khối, chiều cao cây, diện tích lá qua ảnh chụp RGB**

PlantCV là mã nguồn mở bằng python dùng để phân tích HTP do Gehan & cs. (2017) phát triển. Đây là công cụ chuẩn hóa, nhận dạng hình thái học, phân tách mục tiêu chính xác và có khả năng tái lập chu trình phân tích, trích xuất thông tin từ quá trình xử lý hình ảnh. Các tính trạng mà PlantCV có thể trích xuất được bao gồm chiều cao, số lá, năng suất sinh khối, màu sắc, kiến trúc thực vật (từ hình chụp kỹ thuật số RGB). Nhiều các tính trạng khác cũng có thể được trích xuất từ Plant CV nếu hình chụp bằng cảm biến siêu phổ và hồng ngoại. Công cụ Plant CV đã được áp dụng thành công trên nhiều cây trồng khác nhau như ngô (Enders & cs., 2019), khoai lang (Rosero & cs., 2019) và trên một số loại cây trồng khác.

### **5.3. Đánh giá kiểu hình bắp ngô, bông lúa**

Cấu trúc bắp ngô là tính trạng quan trọng trong chọn tạo giống ngô. Phương pháp đánh giá kiểu hình bắp ngô qua ảnh được Miller & cs. (2017) và Makanza & cs. (2018) đề xuất. Dựa trên bức ảnh chụp bắp bằng máy ảnh thông thường được cố định khoảng cách chụp kết hợp thuật toán xử lý hình ảnh nhanh chóng xây dựng từ phần mềm xử lý ảnh miễn phí ImageJ chuyển dữ liệu hàng nghìn hình ảnh bắp sang tính trạng chiều dài, chiều rộng, đặc điểm bắp, với độ chính xác rất cao.

Phương pháp đánh giá cấu trúc bông lúa với công cụ P-TRAP được Al-Tam & cs. (2013) phát triển. Crowell & cs. (2014) hoàn thiện độ chính xác và công bố toàn bộ chu trình trích xuất dữ liệu từ cấu trúc bông bằng ngôn ngữ python đặt tên là PANorama. Crowell & cs. (2016) kết hợp phương pháp PANorama đánh giá 49 kiểu hình bông lúa từ 242 nguồn vật liệu lúa nhiệt đới và phương pháp phân tích liên kết toàn hệ gen (GWAS) giải mã các gen quy định cấu trúc bông. Tuy nhiên yêu cầu của phương pháp cần phân tách riêng rẽ cấu trúc bông, cố định bông, tránh chồng xếp để giảm sai số khi chuyển đổi hình ảnh.

### **5.4. Đánh giá số lượng, hình dạng, cấu trúc hạt**

Hình dạng và kích thước hạt là tính trạng quan trọng trong chọn tạo giống, ảnh hưởng tới mẫu mã, phẩm chất và năng suất hạt. Đo đếm bằng phương pháp truyền thống cho thấy tốn kém thời gian và công sức. Tanabata & cs. (2012) đã phát triển phần mềm SmartGrain cho phép chuyển đổi dữ liệu từ hình ảnh quét từ máy scan hoặc máy ảnh với độ phân giải cao sang 8 tính trạng quan trọng liên quan đến hình dạng và kích thước hạt lúa. Phương pháp cũng cho phép xử lý nhiều bức ảnh chứa trong từng thư mục trong thời gian ngắn. Phương pháp có hạn chế là không nhận dạng được hạt xếp chồng lên nhau. Phương pháp của Tanabata & cs. (2012) còn rất hiệu quả đối với các hạt của cây trồng khác như đậu tương, đậu đỗ, ngô, lúa mì, dưa chuột, dưa lưới.

### **5.5. Dự đoán diện tích, năng suất qua ảnh chụp từ flycam**

Dự đoán diện tích, năng suất lúa qua ảnh được tiến hành trên mô hình bản đồ Việt Nam xây dựng từ cây lúa do giảng viên bộ môn Di truyền và Chọn giống cây trồng và sinh viên Câu lạc bộ Chọn giống cây trồng thuộc Khoa Nông học, Học viện Nông nghiệp Việt Nam thực hiện trong vụ xuân năm 2020 (Hình 5). Giống lúa sử dụng trong mô hình là giống lúa thuần chất lượng cao ĐH12 do Trần Văn Quang và cộng sự chọn tạo. Ảnh được chụp từ flycam, với logo Học viện Nông nghiệp Việt Nam là thước đo cho hình ảnh có đường kính 110cm.

Các bước tiến hành xử lý ảnh bằng phần mềm ImageJ bao gồm tách phần ảnh cần dự đoán (Image segmentation), chuyển sang định dạng 8bit, tạo Binary (Make Binary) đặt thước cho hình ảnh (Set measurement scale) và phân tích thành phần (Analyze Particles) (Hình 5). Tiến hành lấy mẫu 1m<sup>2</sup> giống lúa ĐH12 cho năng suất 0,7kg. Kết quả dự đoán cho thấy diện tích khu vực bản đồ Việt Nam bao gồm hai quần đảo Hoàng Sa, Trường Sa, đảo Phú Quốc và chữ ĐH12 có tổng là 150,77m<sup>2</sup>. Năng suất dự kiến thu hoạch 105,54kg trong khi năng suất thực thu là 92,5kg. Phương pháp đơn giản, hiệu quả,

có độ chính xác rất cao giúp các nhà khoa học định lượng chính xác diện tích và năng suất cây trồng không chỉ trên lúa mà còn có thể áp dụng trên các loại cây trồng khác.

### 5.6. Chi phí xây dựng và vận hành hệ thống HTP

Nghiên cứu về chi phí đầu tư và vận hành các hệ thống HTP trong nhà kính và ngoài đồng ruộng bước đầu đã được Reynolds & cs. (2019) nghiên cứu, nhóm tác giả tổng hợp tại bảng 4.

Cỡ mẫu giả định là 1.700 ô thí nghiệm/cây (284 nguồn gen cây trồng  $\times$  2 thí nghiệm  $\times$  3 lần nhắc lại) được tiến hành thu thập dữ liệu trong 10 ngày đối với thí nghiệm ngoài đồng ruộng và 90 ngày đối với thí nghiệm trong nhà kính. Mỗi ô nhỏ thí nghiệm có diện tích 4-10m<sup>2</sup>. Hệ thống máy kéo tự động, thiết bị bay không người lái và hệ thống HTP trong nhà kính đều được trang bị cảm biến cơ bản như RGB và LIDAR. Chi phí xây dựng nhà kính và khu thí nghiệm đồng ruộng không được tính trong dự toán này. Hai kịch bản đã được đưa ra để tính chi phí gồm kịch bản 1: việc sử dụng các thiết bị bị giới hạn bởi số lượng thiết bị hoặc nhân sự và kịch bản 2: giới hạn bởi nhu cầu phân tích của người dùng hoặc doanh nghiệp tư nhân. Thí nghiệm trong nhà kính được tính theo kịch bản 1. Kết quả cho

thấy trong cả hai kịch bản thì thấy tổng chi phí cho mỗi ô thí nghiệm nhỏ ngoài đồng ruộng hoặc cho mỗi chậu thí nghiệm trong nhà kính là tương đương nhau. Như vậy, chi phí cho các phương tiện mang theo cảm biến chỉ chiếm khoảng 5-26% tổng chi phí đầu tư. Chi phí tương đối này phụ thuộc vào bối cảnh, đặc biệt là giá công lao động, nhu cầu cỡ mẫu, số lượng kiểu hình cần đánh giá và thời gian đánh giá. Chi phí phân tích dữ liệu chiếm khoảng 10-20% tổng chi phí nếu phương pháp phân tích đã được phát triển. Các thiết bị đánh giá kiểu hình cầm tay giá rẻ thích hợp cho các phép đo đếm các tính trạng cây trồng nhanh chóng và không có tính thường xuyên, liên tục. Trong các thí nghiệm phân tích di truyền hoặc đánh giá các đặc điểm nông học, các chi phí chính thường phát sinh từ việc quản lý cây trồng và nhân lực vận hành hệ thống. Nhìn chung, nếu đầu tư ban đầu vào các cảm biến, phương tiện và phát triển phương pháp phân tích, trích xuất dữ liệu thì sẽ thu được chất lượng kiểu hình cao hơn với chi phí vận hành thấp hơn. Như vậy, hạn chế lớn nhất khi xây dựng hệ thống HTP tại Việt Nam hiện nay là kinh phí đầu tư hệ thống, thiếu nhân lực liên ngành với chuyên môn tay nghề cao để phân tích cơ sở dữ liệu hình ảnh thu thập được từ các hệ thống HTP.



Hình 5. Các bước xử lý hình ảnh và phân tách khu vực trồng lúa cần dự đoán

**Bảng 4. Chi phí xây dựng và vận hành hệ thống HTP quy mô nhà lưới và đồng ruộng**

Các hệ thống HTP	Cơ mẫu và thời gian đo đếm	Thời gian sử dụng thiết bị (năm)	Đơn vị: 1000 USD									
			Đầu tư hệ thống	Vận hành & đào tạo nhân lực	Bảo trì hàng năm	Chăm sóc thí nghiệm (I)	Thu thập hình ảnh (II)	Phân tích hình ảnh (III)	Phân tích tình trạng (IV)	Lưu trữ dữ liệu trong 10 năm (V)	Tổng (I + II + III + IV + V)	Chiếm % đầu tư
Thí nghiệm đánh giá ngoài đồng ruộng tính theo kịch bản 1: Giới hạn bởi sự sẵn có của thiết bị và nhân sự												
Máy kéo tự động	1.700 ô thí nghiệm nhỏ trong 10 ngày	20	430,0	19,6	15,0	68,0	17,4	3,5	5,3	1,5	95,7	18,2
Thiết bị cầm tay		15	50,0	15,6	3,0	68,0	16,7	5,3	7,1	0,7	97,8	17,1
Máy bay không người lái (UAV)		2	10,0	24,5	2,0	68,0	4,9	7,1	10,6	0,2	90,8	5,4
Thí nghiệm đánh giá ngoài đồng ruộng tính theo kịch bản 2: Giới hạn bởi nhu cầu người dùng												
Máy kéo tự động	1.700 ô thí nghiệm nhỏ trong 10 ngày	20	430,0	12,9	15,0	68,0	28,4	3,5	5,3	1,5	106,7	26,6
Thiết bị cầm tay		15	50,0	15,6	3,0	68,0	16,7	5,3	7,1	0,7	97,8	17,1
Máy bay không người lái (UAV)		2	10,0	17,0	2,0	68,0	16,6	7,1	10,6	0,2	102,5	16,2
Thí nghiệm trong nhà kính tính theo kịch bản 1: Giới hạn bởi sự sẵn có của thiết bị và nhân sự												
Hệ thống HTP trong nhà kính	1.700 cây trong 90 ngày	15	1000,0	103,6	15,0	71,2	9,0	1,8	10,6	2,6	95,2	9,5

## 6. KẾT LUẬN

Việc đánh giá kiểu hình vẫn là nền tảng của quá trình chọn giống cây trồng. Những tiến bộ trong công nghệ hình ảnh, cảm biến và khoa học máy tính đã thúc đẩy sự phát triển của các phương pháp HTP. Sự khác biệt lớn của phương pháp HTP là có thể đánh giá kiểu hình cây trồng theo thời gian, nhiều chiều và rất đa dạng bao gồm rễ, thân, lá, hạt, cấu trúc bông, bắp cùng nhiều tính trạng hình học khác mà phương pháp truyền thống không thể làm được. Mặc dù có rất nhiều ưu điểm, nhưng dữ liệu tạo ra từ các phương pháp HTP ngày càng lớn đã mang đến bài toán dữ liệu lớn. Để giải bài toán về dữ liệu lớn trong khi xử lý dữ liệu HTP, việc xây dựng một hệ sinh thái tính toán tiên tiến được hỗ trợ bởi các công cụ học máy, học sâu cho phép lưu trữ đồng thời xử lý và phân tích dữ liệu là yếu tố then chốt để đưa ra các suy luận có ý nghĩa từ các tập dữ liệu HTP đa chiều.

Với những triển vọng từ phương pháp đánh giá kiểu hình hiệu năng cao trên cây trồng - HTP, Việt Nam hoàn toàn có thể áp dụng ngay các thành tựu hiện tại từ phương pháp này trong việc đánh giá kiểu hình qua tất cả các giai đoạn phát triển của cây trồng. Trong thập niên tới, để phát triển phương pháp một cách hệ thống cần có sự kết hợp nghiên cứu liên ngành bao gồm sinh lý thực vật, chọn giống cây trồng, công nghệ sinh học, ngành công nghệ thông tin, khoa học máy tính, xử lý ảnh, viễn thám, thống kê, tin sinh học. Cùng với đó, cần sớm xây dựng cơ sở dữ liệu lớn, có thể đặt tên VNAgribigdata không chỉ phục vụ cho việc thu thập, lưu trữ, phân tích, xử lý dữ liệu HTP mà còn kết hợp với các thông tin bộ gen cây trồng để chuyển trạng thái chọn lọc cây trồng sang thiết kế cây trồng theo từng vùng sinh thái trong tương lai.

## LỜI CẢM ƠN

Nghiên cứu này nằm trong Đề tài khoa học và công nghệ trọng điểm cấp Học viện 2021: “Nghiên cứu mở đầu đánh giá kiểu hình qua ảnh, xây dựng dữ liệu lớn và ứng dụng trong chọn giống cây trồng” mã số: T2021-41-15TD.

Nhóm tác giả trân trọng cảm ơn Học viện Nông nghiệp Việt Nam đã tài trợ kinh phí thực hiện nghiên cứu này.

Nhóm tác giả cũng trân trọng cảm ơn hai chuyên gia phản biện độc lập và Ban biên tập đã có những góp ý vô cùng quý báu để hoàn thiện nghiên cứu tổng quan này.

## TÀI LIỆU THAM KHẢO

- Al-Tam F., Adam H., Anjos A.d., Lorieux M., Larmande P., Ghesquière A., Jouannic S. & Shahbazkia H. R. (2013). P-TRAP: a Panicle Trait Phenotyping tool. *BMC Plant Biology*. 13(1): 122.
- Araus J.L. & Cairns J.E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*. 19(1): 52-61.
- Araus J.L., Kefauver S.C., Zaman-Allah M., Olsen M.S. & Cairns J.E. (2018). Translating high-throughput phenotyping into genetic gain. *Trends in Plant Science*. 23(5): 451-466.
- Behjati S. & Tarpey P.S. (2013). What is next generation sequencing? *Archives of Disease in Childhood: Education & Practice*. 98(6): 236-238.
- Beres B.L., Hatfield J.L., Kirkegaard J.A., Eigenbrode S.D., Pan W.L., Lollato R.P., Hunt J.R., Strydhorst S., Porker K., Lyon D., Ransom J. & Wiersma J. (2020). Toward a Better Understanding of Genotype  $\times$  Environment  $\times$  Management Interactions - A Global Wheat Initiative Agronomic Research Strategy. *Frontiers in Plant Science*. 11(828).
- Burton A.L., Williams M., Lynch J.P. & Brown K.M. (2012). RootScan: Software for high-throughput analysis of root anatomical traits. *Plant and soil*. 357(1): 189-203.
- Carroll A.A., Clarke J., Fahlgren N., Gehan M.A., Lawrence-Dill C.J. & Lorence A. (2019). NAPPN: Who We Are, Where We Are Going, and Why You Should Join Us! *The Plant Phenome Journal*. 2(1).
- Cobb J.N., Declerck G., Greenberg A., Clark R. & Mccouch S. (2013). Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*. 126(4): 867-887.
- Crowell S., Falcão A.X., Shah A., Wilson Z., Greenberg A.J. & Mccouch S.R. (2014). High-Resolution Inflorescence Phenotyping Using a Novel Image-Analysis Pipeline, PANorama. *Plant Physiology*. 165(2): 479-495.

- Crowell S., Korniliev P., Falcão A., Ismail A., Gregorio G., Mezey J. & Mccouch S. (2016). Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nature Communications*. 7(1): 10527.
- Danilevicz M.F., Bayer P.E., Nestor B.J., Bennamoun M. & Edwards D. (2021). Resources for image-based high-throughput phenotyping in crops and data sharing challenges. *Plant Physiology*. 10.1093/plphys/kiab301.
- Deery D.M., Rebetzke G.J., Jimenez-Berni J.A., James R.A., Condon A.G., Bovill W.D., Hutchinson P., Scarrow J., Davy R. & Furbank R.T. (2016). Methodology for High-Throughput Field Phenotyping of Canopy Temperature Using Airborne Thermography. *Frontiers in Plant Science*. 7(1808).
- Enders T.A., St. Dennis S., Oakland J., Callen S.T., Gehan M.A., Miller N.D., Spalding E.P., Springer N.M. & Hirsch C.D. (2019). Classifying cold-stress responses of inbred maize seedlings using RGB imaging. *Plant Direct*. 3(1): e00104.
- Falk K.G., Jubery T.Z., Mirnezami S.V., Parmley K.A., Sarkar S., Singh A., Ganapathysubramanian B. & Singh A. K. (2020a). Computer vision and machine learning enabled soybean root phenotyping pipeline. *Plant Methods*. 16(1): 5.
- Falk K.G., Jubery T.Z., O'rourke J.A., Singh A., Sarkar S., Ganapathysubramanian B. & Singh A.K. (2020b). Soybean Root System Architecture Trait Study through Genotypic, Phenotypic, and Shape-Based Clusters. *Plant Phenomics*. 2020: 1925495.
- Fiorani F. & Schurr U. (2013). Future Scenarios for Plant Phenotyping. *Annual Review of Plant Biology*. 64(1): 267-291.
- Furbank R.T. & Tester M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*. 16(12): 635-644.
- Gehan M.A., Fahlgren N., Abbasi A., Berry J.C., Callen S.T., Chavez L., Doust A.N., Feldman M.J., Gilbert K.B., Hodge J.G., Hoyer J.S., Lin A., Liu S., Lizárraga C., Lorence A., Miller M., Platon E., Tessman M. & Sax T. (2017). PlantCV v2: Image analysis software for high-throughput plant phenotyping. *Peer J*. 5: e4088.
- Hartmann A., Czauderna T., Hoffmann R., Stein N. & Schreiber F. (2011). HTPheno: An image analysis pipeline for high-throughput plant phenotyping. *BMC Bioinformatics*. 12(1): 148.
- Hickey L.T., Hafeez N.A., Robinson H., Jackson S.A., Leal-Bertioli S.C.M., Tester M., Gao C., Godwin I.D., Hayes B.J. & Wulff B.B.H. (2019). Breeding crops to feed 10 billion. *Nature Biotechnology*. 37(7): 744-754.
- Johannsen W. (1911). The genotype conception of heredity. *International Journal of Epidemiology*. 43(4): 989-1000.
- Klukas C., Chen D. & Pape J.M. (2014). Integrated Analysis Platform: An Open-Source Information System for High-Throughput Plant Phenotyping. *Plant Physiol*. 165(2): 506-518.
- Knecht A.C., Campbell M.T., Caprez A., Swanson D.R. & Walia H. (2016). Image Harvest: an open-source platform for high-throughput plant image processing and analysis. *Journal of Experimental Botany*. 67(11): 3587-3599.
- Lobet G., Pagès L. & Draye X. (2011). A Novel Image-Analysis Toolbox Enabling Quantitative Analysis of Root System Architecture *Plant Physiology*. 157(1): 29-39.
- Lynch J.P. (2019). Root phenotypes for improved nutrient capture: an underexploited opportunity for global agriculture. *New Phytologist*. 223(2): 548-564.
- Makanza R., Zaman-Allah M., Cairns J. E., Eyre J., Burgueño J., Pacheco Á., Diepenbrock C., Magorokosho C., Tarekegne A., Olsen M. & Prasanna B. M. (2018). High-throughput method for ear phenotyping and kernel weight estimation in maize using ear digital imaging. *Plant Methods*. 14(1): 49.
- Miller N.D., Haase N. J., Lee J., Kaeppler S.M., De Leon N. & Spalding E.P. (2017). A robust, high-throughput method for computing maize ear, cob, and kernel attributes automatically from images. *The Plant Journal*. 89(1): 169-178.
- Nguyen T.D. (2020). High-throughput phenotyping of rice genotypes for nitrogen use efficiency. ICAR-Indian Agricultural Research Institute, New Delhi. Master Thesis: T-10424.
- Pautasso M. (2013). Ten simple rules for writing a literature review. *PLOS Computational Biology*. 9(7): e1003149.
- Pound M.P., French A.P., Murchie E.H. & Pridmore T.P. (2014). Automated Recovery of Three-Dimensional Models of Plant Shoots from Multiple Color Images *Plant Physiology*. 166(4): 1688-1698.
- Reynolds D., Baret F., Welcker C., Bostrom A., Ball J., Cellini F., Lorence A., Chawade A., Khafif M., Noshita K., Mueller-Linow M., Zhou J. & Tardieu F. (2019). What is cost-efficient phenotyping? Optimizing costs for different scenarios. *Plant Science*. 282: 14-22.
- Rosero A., Granda L., Pérez J.L., Rosero D., Burgos-Paz W., Martínez R., Morelo J., Pastrana I., Burbano E. & Morales A. (2019). Morphometric and colourimetric tools to dissect morphological diversity: an application in sweet potato [*Ipomoea batatas* (L.) Lam.]. *Genetic Resources and Crop Evolution*. 66(6): 1257-1278.

- Schork N.J. (1997). Genetics of Complex Disease. *American Journal of Respiratory and Critical Care Medicine*. 156(4): S103-S109.
- Seethepalli A., Guo H., Liu X., Griffiths M., Almtarfi H., Li Z., Liu S., Zare A., Fritschi F. B., Blancaflor E.B., Ma X.F. & York L.M. (2020). RhizoVision Crown: An Integrated Hardware and Software Platform for Root Crown Phenotyping. *Plant Phenomics*. p. 3074916.
- Shi Y., Thomasson J.A., Murray S.C., Pugh N.A., Rooney W.L., Shafian S., Rajan N., Rouze G., Morgan C.L.S., Neely H.L., Rana A., Bagavathiannan M.V., Henrickson J., Bowden E., Valasek J., Olsenholler J., Bishop M. P., Sheridan R., Putman E.B., Popescu S., Burks T., Cope D., Ibrahim A., Mccutchen B.F., Baltensperger D.D., Avant R.V., Jr., Vidrine M. & Yang C. (2016). Unmanned Aerial Vehicles for High-Throughput Phenotyping and Agronomic Research. *PLOS ONE*. 11(7): e0159781.
- Tanabata T., Shibaya T., Hori K., Ebana K. & Yano M. (2012). SmartGrain: High-Throughput Phenotyping Software for Measuring Seed Shape through Image Analysis *Plant Physiology*. 160(4): 1871-1880.
- Tanger P., Klassen S., Mojica J.P., Lovell J.T., Moyers B.T., Baraoidan M., Naredo M.E.B., McNally K.L., Poland J., Bush D.R., Leung H., Leach J.E. & McKay J.K. (2017). Field-based high throughput phenotyping rapidly identifies genomic regions controlling yield components in rice. *Scientific Reports*. 7: 42839.
- Tardieu F., Cabrera-Bosquet L., Pridmore T. & Bennett M. (2017). Plant Phenomics, From Sensors to Knowledge. *Current Biology*. 27(15): R770-R783.
- Tuberosa R. (2012). Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology*. 3(347).
- Vadez V., Kholová J., Hummel G., Zhokhavets U., Gupta S.K. & Hash C.T. (2015). LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget. *Journal of Experimental Botany*. 66(18): 5581-5593.
- Varshney R.K., Bohra A., Roorkiwal M., Barmukh R., Cowling W.A., Chitkineni A., Lam H.M., Hickey L.T., Croser J.S., Bayer P.E., Edwards D., Crossa J., Weckwerth W., Millar H., Kumar A., Bevan M.W. & Siddique K.H.M. (2021). Fast-forward breeding for a food-secure world. *Trends in Genetics*. <https://doi.org/10.1016/j.tig.2021.08.002>.
- Walter A., Liebisch F. & Hund A. (2015). Plant phenotyping: from bean weighing to image analysis. *Plant Methods*. 11(1): 14.
- Whan A.P., Smith A.B., Cavanagh C.R., Ral J.P.F., Shaw L.M., Howitt C.A. & Bischof L. (2014). GrainScan: a low cost, fast method for grain size and colour measurements. *Plant Methods*. 10(1): 23.
- Yang W., Feng H., Zhang X., Zhang J., Doonan J.H., Batchelor W.D., Xiong L. & Yan J. (2020). Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives. *Molecular Plant*. 13(2): 187-214.
- Yang W., Guo Z., Huang C., Duan L., Chen G., Jiang N., Fang W., Feng H., Xie W., Lian X., Wang G., Luo Q., Zhang Q., Liu Q. & Xiong L. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature Communications*. 5: 5087.
- Zhao C., Zhang Y., Du J., Guo X., Wen W., Gu S., Wang J. & Fan J. (2019). Crop Phenomics: Current Status and Perspectives. *Frontiers in Plant Science*. 10(714).