

THÔNG TIN LUẬN ÁN TIẾN SĨ

Họ và tên NCS: **TRẦN THỊ KIM NGÀ**
Đề tài nghiên cứu: **NGHIÊN CỨU NÂNG CAO ĐỘ CHÍNH XÁC PHÂN LOẠI
GIỐNG LÚA THÔNG QUA ẢNH HẠT LÚA**
Chuyên ngành: **KỸ THUẬT ĐIỆN TỬ**
Mã số chuyên ngành: **62520203**
Tập thể hướng dẫn: **PGS. TS. ĐỖ HỒNG TUẤN; TS. VLADIMIR Y. MARIANO**
Cơ sở đào tạo: **Trường Đại học Bách Khoa, Đại học Quốc gia Thành phố Hồ Chí Minh.**

Mục tiêu của Luận án

Mỗi giống lúa sẽ thích hợp với điều kiện gieo trồng như khí hậu, thổ nhưỡng nhất định. Nên việc chọn giống lúa phù hợp với điều kiện sinh trưởng là một trong những bước quan trọng góp phần nâng cao năng suất và chất lượng của vụ mùa. Tuy nhiên, nhiều giống lúa mới được tạo ra để thích ứng với các thay đổi như khí hậu, sâu bọ và dịch bệnh, v.v nên chúng trở nên ngày càng đa dạng. Bên cạnh các giống lúa với quan sát bên ngoài của hạt giống có sự khác nhau rõ ràng, thì có một số giống lúa khá giống nhau về màu sắc, hình dạng, bề mặt, v. v của hạt lúa nên dễ gây nhầm lẫn với nhau. Điều này dẫn đến dễ chọn nhầm giống lúa và vì vậy năng suất và chất lượng của vụ mùa không cao. Do đó, việc nâng cao độ chính xác phân loại cho các loại giống lúa có quan sát bên ngoài tương tự nhau trở nên vô cùng cần thiết.

Chính vì vậy, mục tiêu của nghiên cứu này nhằm nâng cao độ chính xác phân loại của 17 giống lúa có quan sát bên ngoài hạt lúa khá giống nhau về màu sắc được trồng ở Việt Nam. Phương pháp trích xuất đặc trưng của ảnh lúa kết hợp với phương pháp máy học được thực hiện. Tập mô tả gồm đặc trưng về màu sắc, hình thái học và màu sắc kết hợp với kỹ thuật tối ưu bầy đàn nhị phân (BPSO) và máy vector hỗ trợ (SVM), gọi là BPSO+SVM, để đánh giá và lựa chọn tập đặc trưng hiệu quả nhằm nâng cao độ chính xác phân loại. Bên cạnh phương pháp phân tích bề mặt dựa trên ma trận đồng hiện mức xám, phương pháp phân tích bề mặt dựa trên mẫu bộ ba cục bộ (LTP) cũng được đề xuất để kết hợp với SVM và mạng nơ ron nhân tạo (ANN). Để nâng cao độ chính xác phân loại, hai mô hình mạng nơ ron tích chập được hiệu chỉnh từ VGG16 và ResNet50 được kết hợp cùng với cơ sở dữ liệu ảnh để phân loại 17 giống lúa trên.

Những đóng góp chính của Luận án

Những đóng góp chính của nghiên cứu được tóm tắt như sau:

- Nghiên cứu đã xây dựng ba cơ tập dữ liệu để phân loại 17 giống lúa. Tập dữ liệu thứ nhất được gọi là tập đặc trưng tổng cộng, bao gồm 5 tập con có tên là màu cơ bản, màu phân

cum, hình thái học, bề mặt thống kê và ma trận đồng hiện mức xám (GLCM). Tập dữ liệu thứ hai được gọi là tập đặc trưng mẫu bộ ba cục bộ cải tiến mở rộng (extended improved local ternary pattern – extended ILTP). Và tập dữ liệu thứ ba là cơ sở dữ liệu ảnh, được xây dựng để kết hợp với mạng nơ ron tích chập (CNN). Cơ sở dữ liệu này gồm 7 tập ảnh lúa với số lượng ảnh trong tập huấn luyện khác nhau.

- Để giảm số đặc trưng sử dụng và cải thiện độ chính xác phân loại, phương pháp đề xuất BPSO+SVM, được áp dụng với tập thứ nhất để phân loại 17 giống lúa. Kết quả nhận được cho thấy phương pháp đề xuất BPSO+SVM đạt độ chính xác phân loại cao hơn so với khi sử dụng chỉ với SVM, và số đặc trưng được sử dụng chỉ bằng 39% số lượng đặc trưng của tập tổng cộng. Hơn nữa, phương pháp đề xuất BPSO+SVM cũng có thể mở rộng cho những bộ phân loại khác như Naïve Bayes và Random Forest.

- Tập dữ liệu thứ hai được đề xuất là ILTP mở rộng, được kết hợp với SVM và ANN để phân loại 17 giống lúa. Để cải thiện độ chính xác phân loại, bán kính vùng cục bộ và giá trị ngưỡng của vùng cục bộ cũng được xem xét trong mô tả đề xuất. Độ chính xác phân loại 17 giống lúa được cải thiện khi kết hợp tập đặc trưng ILTP mở rộng với SVM.

- Hai mô hình mạng nơ ron tích chập được đề xuất được gọi là VGG16 hiệu chỉnh và ResNet50 hiệu chỉnh được xây dựng dựa trên các mô hình VGG16 và ResNet50. Các thí nghiệm được tiến hành để đánh giá và so sánh hiệu suất khi thay đổi số lượng ảnh trong tập huấn luyện, cũng như so sánh độ chính xác phân loại của hai mô hình trước và sau hiệu chỉnh. Kết quả cho thấy các mô hình hiệu chỉnh đạt hiệu quả phân loại tốt hơn so với mô hình chưa hiệu chỉnh. Như vậy, hai mô hình hiệu chỉnh VGG16 và ResNet50 có thể nâng cao đáng kể độ chính xác phân loại 17 giống lúa.

- Cuối cùng, nghiên cứu cũng được mở rộng để đánh giá khả năng phân loại bền vững của các mô hình đối với các trường hợp: (1) hạt giống lúa sau thời gian dài lưu trữ ở nhiệt độ thường, (2) trường hợp hạt giống lúa bị đổi màu do ẩm, mốc, (3) mở rộng phân loại đối với các giống lúa khác, (4) và cuối cùng mở rộng phân loại với 34 giống lúa. Các kết quả cho thấy các mô hình này có thể mở rộng để phân loại cho các mẫu lúa sau thời gian dài lưu trữ, các mẫu lúa bị thay đổi màu sắc do ẩm, mốc, có thể mở rộng phân loại hiệu quả với các giống lúa khác cũng như có thể mở rộng để phân loại nhiều giống lúa hơn.

Thành phố Hồ Chí Minh, ngàytháng.....năm.....

Tập thể hướng dẫn

Nghiên cứu sinh

PGS. TS. ĐỖ HỒNG TUẤN TS. VLADIMIR Y. MARIANO TRẦN THỊ KIM NGÀ