

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC BÁCH KHOA
KHOA ĐIỆN TỬ - VIỄN THÔNG

BÁO CÁO KẾT THÚC HỌC PHẦN
KỸ THUẬT NHẬN DẠNG

Đề tài: Nhận dạng các loại thảo mộc và gia vị

LỚP HP: 18.40

NHÓM: 6A

GVHD: HOÀNG LÊ UYÊN THỰC

SVTH: Đặng Ngọc Quốc Bảo 18DT2

Lê Trí Đô 18DT2

Lê Văn Mẫn 18DT2

Nguyễn Anh Quân 18DT2

Lê Tuấn Anh 18DT2

Đà Nẵng, 12/2021

Lời cảm ơn

*Chúng em xin gửi lời cảm ơn chân thành đến giảng viên hướng dẫn **TS. Hoàng Lê Uyên Thực** đã tạo điều kiện thuận lợi cho chúng em học tập và hoàn thành đề tài này. Chúng em xin bày tỏ lòng biết ơn sâu sắc đến cô đã dày công truyền đạt kiến thức và hướng dẫn chúng em trong quá trình làm bài. Chúng em đã cố gắng vận dụng những kiến thức đã học được trong kỳ học để hoàn thành dự án. Nhưng do không có nhiều kinh nghiệm thực tiễn, nên khó tránh khỏi những thiếu sót trong quá trình nghiên cứu và trình bày. Rất mong sự góp ý của quý thầy cô để dự án của nhóm chúng em được hoàn thiện hơn.*

Một lần nữa, em xin trân trọng cảm ơn cô đã giúp đỡ chúng em trong quá trình thực hiện dự án này. Chúc cô luôn mạnh khỏe và thành công trong cuộc sống.






Xin trân trọng cảm ơn!

MỞ ĐẦU

Hiện nay, cùng với sự phát triển của khoa học kỹ thuật việc phát triển con người là mục tiêu cao cả nhất của toàn nhân loại. Con người ngày càng tiến bộ trong lĩnh vực khoa học, phát triển ra được những ứng dụng hữu ích cho đời sống chúng ta. Các hệ thống nhận dạng được ra đời và phát triển với độ tin cậy ngày càng cao. Với cách tiếp cận đối tượng nhận dạng theo phương pháp này, chúng ta có thể thu nhận được nhiều thông tin từ đối tượng hơn mà không cần tác động nhiều đến đối tượng mà vẫn đảm bảo chính xác, an toàn, thân thiện.

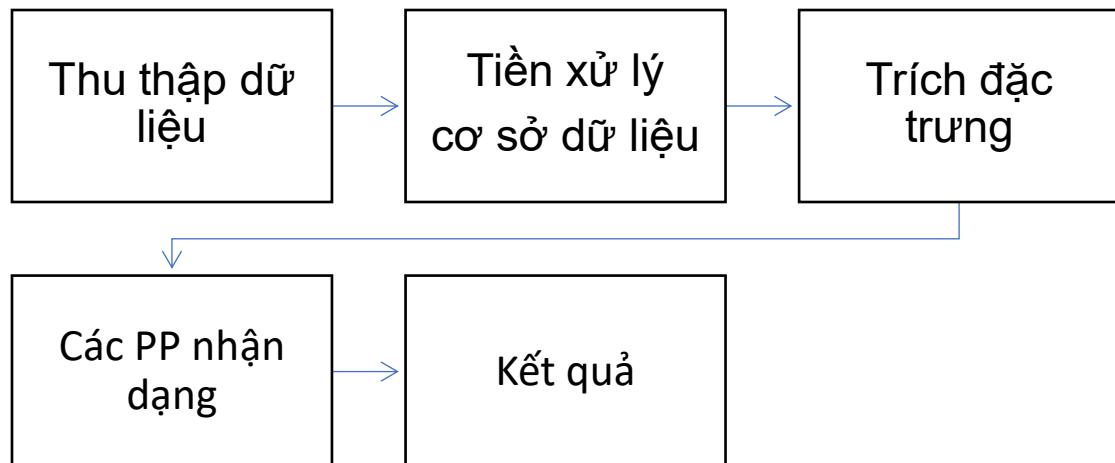
Trong đời sống của chúng ta, các loại thảo mộc và gia vị là những thứ không thể thiếu trong mỗi bữa ăn, nó giúp chúng ta ăn ngon hơn hay chỉ đơn giản là trang trí đĩa thức ăn cho thêm phần đặc sắc. Những điều tưởng như là nhỏ bé nhưng đó chính là sợi dây vô hình kết nối chúng ta lại. Việc xây dựng mô hình nhận dạng giúp cho trẻ nhỏ và mọi người tiếp cận với các loại củ, quả này một cách nhanh chóng, phân biệt tránh nhầm lẫn giữa các loại. Sử dụng trong các dây chuyền phân loại sản phẩm trong nhà máy, siêu thị, ...

Mục tiêu của nhóm mình là hướng đến xây dựng mô hình nhận dạng mẫu các loại thảo mộc và gia vị có ngay trong nhà bếp như gừng, ớt, tỏi, hành, sả,.. với độ chính xác cao, nâng cao kỹ năng lập trình, trích được nhiều đặc trưng, ảnh thu thập chất lượng có chất lượng tốt, cải thiện kỹ năng làm việc nhóm, đưa Machine Learning tiếp cận với đời sống ...

BẢNG PHÂN CÔNG NHIỆM VỤ				
Họ và tên	Lớp	MSSV	Nhiệm vụ	Chữ ký
Đặng Ngọc Quốc Bảo	18DT2	106180066	Thu thập dữ liệu của ớt, xử lý ảnh, viết báo cáo, lập trình matlab	
Lê Trí Đô	18DT2	106180072	Thu thập dữ liệu của tỏi, xử lý ảnh, viết báo cáo, lập trình matlab	
Lê Văn Mẫn	18DT2	106180094	Thu thập dữ liệu của hành, xử lý ảnh, viết báo cáo lập trình matlab	
Nguyễn Anh Quân	18DT2	106180106	Thu thập dữ liệu của sả, xử lý ảnh, làm slide, lập trình matlab	
Lê Tuấn Anh	18DT2	106180064	Thu thập dữ liệu của gừng, xử lý ảnh, làm slide, lập trình matlab	

I. Cơ sở lý thuyết và phương pháp thực hiện:

Các bước thực hiện

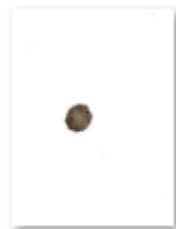
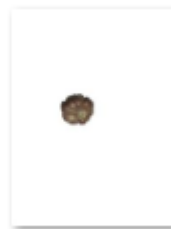


1. Thu thập dữ liệu:

(Mỗi thành viên trong nhóm thu thập ít nhất 100 bức ảnh về các loại đã được phân công)

1.1. Tỏi:

- Số lượng: 100 ảnh
- Một số hình ảnh thu thập được



1.2.Ớt:

- Số lượng: 100 ảnh
- Một số hình ảnh thu thập được



1.3. Gừng:

- Số lượng: 112 ảnh
- Một số hình ảnh thu thập được



1.4. Hành:

- Số lượng: 110 ảnh
- Một số hình ảnh thu thập được



1.5. Sả:

- Số lượng: 100 ảnh
- Một số hình ảnh thu thập được

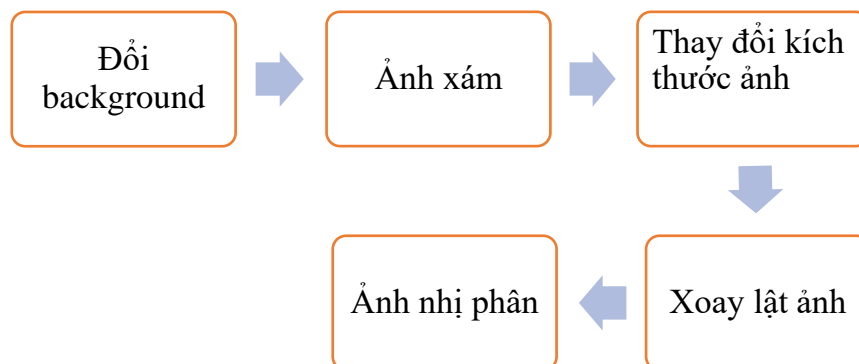


2. Tiền xử lý:

2.1. Ý nghĩa của tiền xử lý:

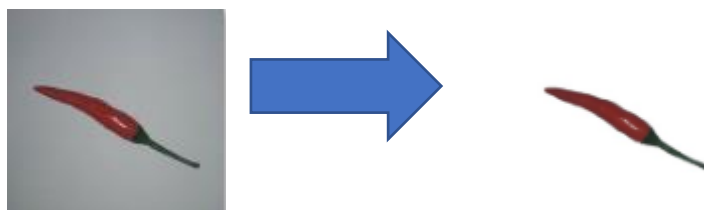
- Tiền xử lý là một bước rất quan trọng để tăng cường chất lượng của dữ liệu, thúc đẩy quá trình trích xuất những thông tin hữu ích cho dữ liệu
- Làm sạch và tổ chức dữ liệu thô phù hợp với việc đào tạo mô hình Machine Learning
- Lọc nhiễu, làm mịn, chuẩn hóa

2.2. Các bước tiền xử lý:



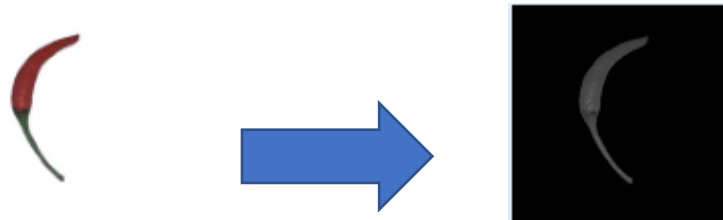
2.2.1. Tách background:

- Sử dụng photoshop để tách background



2.2.2. Ảnh xám:

- Công cụ chính: rgb2gray
- Sử dụng phần mềm matlab
- Ảnh sau khi qua xử lý như sau

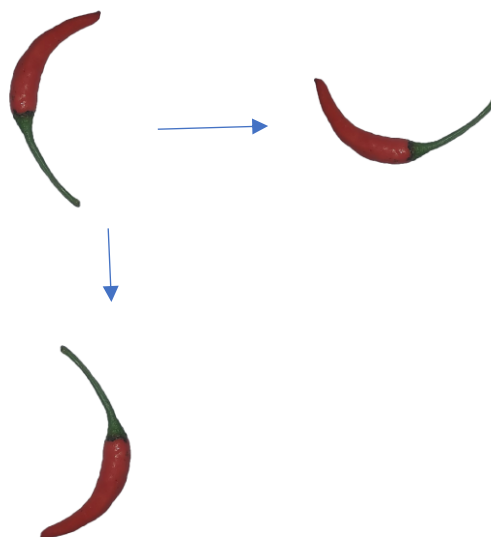


2.2.3. Thay đổi kích thước ảnh:

- Công cụ chính: imresize
- Sử dụng phần mềm matlab
- Ảnh mới có kích thước 50x50

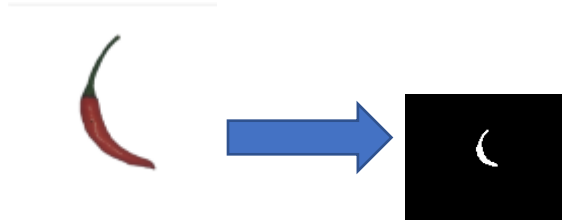
2.2.4. Xoay lật ảnh:

- Sử dụng phần mềm matlab
- Ảnh sau khi qua xử lý như sau



2.2.5. Ảnh nhị phân:

- Sử dụng phần mềm matlab
- Ảnh sau khi qua xử lý như sau



3. Trích đặc trưng:

3.1. Mục đích:

- Để giảm chiều dữ liệu
- Để biến đổi mỗi hình ảnh đầu vào thành một vector đặc trưng đa chiều tương ứng
- Để thực hiện các tác vụ phân loại được xác định trước với độ chính xác đủ mà không cần sử dụng toàn bộ hình ảnh đầu vào
- Trích ra những đặc điểm riêng biệt giữa các đối tượng
- Giảm chiều dữ liệu bằng cách đo những đặc trưng nào đó và dùng các số đo đó để thực hiện phân loại
- Lựa chọn đặc trưng cần tìm hiểu từ chuyên gia, đo độ sáng tối...

3.2. Yêu cầu:

- Các đặc trưng phải trích xuất các đặc điểm phù hợp nhất từ hình ảnh đầu vào

3.3. Hu moment:

Central moments:

$$m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q s(x, y) \quad p, q = 0, 1, 2, 3$$

Central normalized moments:

$$M_{pq} = \frac{m_{pq}}{m_{00}^{\frac{p+q}{2}+1}}$$

Centroid of the image

$$\bar{x} = \frac{\sum_x \sum_y x \cdot s(x, y)}{\sum_x \sum_y s(x, y)}, \bar{y} = \frac{\sum_x \sum_y y \cdot s(x, y)}{\sum_x \sum_y s(x, y)}$$

$$\begin{aligned} S_1 &= M_{20} + M_{02} \\ S_2 &= (M_{20} - M_{02})(M_{20} + M_{02}) + 4M_{11}M_{11} \\ S_3 &= (M_{30} - 3M_{12})^2 + (M_{30} - 3M_{21})^2 \\ S_4 &= (M_{30} + M_{12})^2 + (M_{03} + M_{21})^2 \\ S_5 &= (M_{30} - 3M_{12})(M_{30} + M_{12})[(M_{30} + M_{12})^2 - 3(M_{03} + M_{21})^2] + (3M_{21} - M_{03})(M_{03} + M_{21})[3(M_{30} + M_{12})^2 - (M_{03} + M_{21})^2] \\ S_6 &= (M_{20} - M_{02})[(M_{30} + M_{12})^2 - (M_{03} + M_{21})^2] + 4M_{11}(M_{30} + M_{12})(M_{03} + M_{21}) \\ S_7 &= (3M_{21} - M_{03})(M_{30} + M_{12})[(M_{30} + M_{12})^2 - 3(M_{03} + M_{21})^2] + (M_{30} - 3M_{12})(M_{21} + M_{02})[3(M_{30} + M_{12})^2 - (M_{03} + M_{21})^2] \end{aligned}$$

3.4. Thực hiện trích đặc trưng:

1	0,251742	0,916223	-0,20879	-0,08417	-0,27342	0,261027	0,174895	1
2	0,711033	2,255394	3,182986	1,961326	4,948057	3,752601	-6,19853	1
3	0,456382	1,476823	0,364548	0,580225	0,978995	1,209207	-0,97319	1
4	-0,31406	1,442953	-0,87911	-1,05132	-2,17156	-0,28562	-1,81935	1
5	0,264028	1,866979	-0,12819	1,621648	-0,71617	-1,38942	-1,9951	1
6	0,600617	3,616393	1,694289	1,546878	3,630234	-3,79882	-5,00146	1
7	0,215163	0,668031	-0,31636	0,045584	-0,46666	0,048233	0,04191	1
8	0,285496	2,554619	0,366985	1,260052	-1,39906	-1,96485	-1,74839	1
9	0,742685	2,636136	3,152394	2,098709	-6,02587	5,225741	6,291796	1
10	0,498645	2,010233	1,519121	1,482933	3,329378	2,372805	-3,05494	1
11	0,632845	2,220508	1,732952	1,429312	3,995842	2,83446	3,480087	1
12	0,754727	2,533602	3,97912	2,309326	-7,51032	5,26705	-8,6495	1
13	0,412112	1,565107	1,231315	1,213626	-3,1246	-2,81874	-2,68167	1
14	0,457809	1,637964	1,243723	1,42353	3,131841	2,732285	3,468111	1
15	0,655726	2,748101	2,783608	1,895865	6,714371	4,59685	5,711445	1
16	0,570421	2,519035	1,669745	1,435012	-4,93081	-4,12437	-4,32804	1
17	0,595541	2,032723	1,718007	1,635481	3,905681	3,131891	4,918669	1
18	0,567232	2,135448	1,286254	1,385797	-3,5915	-3,03756	3,177322	1
19	0,54854	2,469184	1,429249	1,452887	-4,76227	-3,77497	-4,47815	1
20	0,672687	3,684661	2,661835	1,751742	-6,03111	5,947024	-5,46846	1
21	0,560364	2,006432	1,026092	1,055413	2,348069	2,099094	-2,45806	1
22	0,727561	2,675594	2,889492	1,954669	6,305769	4,336288	5,80058	1
23	0,51715	3,56383	1,192173	1,694283	-3,56145	-3,74858	-4,50786	1
24	0,754089	3,61718	2,966129	2,051605	-6,92564	-5,3006	-7,27506	1
25	0,464848	1,414694	0,907819	1,360381	1,587665	1,643339	-2,25238	1
26	0,756345	2,912696	3,008961	2,202893	-6,86277	-5,23268	6,355153	1
27	0,503475	2,031573	0,693118	0,767858	1,529922	1,83072	-1,80966	1
28	0,761387	2,870719	3,124494	2,367115	7,596841	-5,60176	-6,98272	1
29	0,670704	2,444538	2,845363	2,457017	-5,88786	-4,74109	-5,84365	1

81	0,197816	0,440657	1,22497	1,186212	2,99568	1,735763	2,893107	2
82	0,261079	0,608141	1,196016	1,388702	2,809699	2,143401	3,169171	2
83	0,316565	0,73593	1,364942	1,67558	4,036306	-3,38591	-5,51777	2
84	0,28531	0,621618	2,388192	0,697351	5,431766	3,259719	5,5981	2
85	0,133624	0,320788	1,06884	0,333239	3,341079	2,681901	-3,24928	2
86	0,127594	0,281754	1,549071	1,499665	2,645603	1,699641	-6,20193	2
87	0,126423	0,317509	1,990818	0,363219	-6,1953	-3,40191	-5,61364	2
88	0,216188	0,711608	0,702759	0,84744	1,970248	1,789718	-2,12006	2
89	0,115437	0,272846	1,425482	0,707512	-4,59455	-2,78664	3,665286	2
90	0,202025	0,704491	0,968576	0,513116	3,018198	2,434986	-3,03033	2
91	0,287501	0,640279	2,477694	2,164833	6,108251	3,856889	-5,58966	2
92	0,197777	0,433948	1,992019	1,46682	-5,16227	-3,15309	-4,55938	2
93	0,236328	0,540303	1,992872	0,653417	4,649137	2,865934	-5,36958	2
94	0,245349	0,793956	0,824862	0,590727	-2,93798	-2,0788	3,40697	2
95	0,200615	0,438778	1,755205	0,639337	3,769631	2,072828	-3,8895	2
96	0,175719	0,444831	0,725157	0,401858	3,189663	4,304581	-2,46024	2
97	0,197816	0,440657	1,22497	1,186212	2,99568	1,735763	2,893107	2
98	0,204315	0,690514	0,658201	0,432845	2,085793	1,79658	-2,31744	2
99	0,300626	0,696473	1,379533	1,956419	-4,01789	-2,69916	3,904539	2
100	0,200615	0,438778	1,755205	0,639337	3,769631	2,072828	-3,8895	2
101	0,1983	0,48269	0,841827	0,641708	2,405687	1,883022	2,663916	2
102	0,290267	0,636779	1,793456	0,686684	3,662962	2,122431	4,385224	2
103	0,125748	0,288616	1,755435	2,60408	-4,8509	-2,78992	5,525953	2
104	0,19331	0,474047	1,28259	0,782471	2,899728	1,782924	-3,87289	2
105	0,200615	0,438778	1,755205	0,639337	3,769631	2,072828	-3,8895	2
106	0,261079	0,608141	1,196016	1,388702	2,809699	2,143401	3,169171	2
107	0,263203	0,605431	2,092226	0,609827	5,088013	3,267832	6,69673	2
108	0,19838	0,674381	0,635844	0,415128	2,027091	1,739093	-2,26973	2
109	0,2221066	0,529099	1,134712	0,750167	-3,47932	-2,39276	3,305568	2
110	0,261079	0,608141	1,196016	1,388702	2,809699	2,143401	3,169171	2

161	0,733708	3,480004	2,920036	1,947193	-7,47285	5,793792	7,577642	3
162	0,601681	1,707281	2,925605	2,261051	7,15516	4,68526	-7,1494	3
163	0,745534	2,833937	2,97687	2,249442	7,864362	6,403778	8,59967	3
164	0,649566	1,929359	3,235838	2,152661	-9,07211	5,961779	-8,61217	3
165	0,682063	1,804333	3,259795	1,554596	-7,9745	-5,08457	-8,09597	3
166	0,706735	2,53891	3,234403	1,978554	7,302883	-5,25519	-7,44674	3
167	0,716545	2,069951	3,754452	2,546452	-8,62422	-5,42031	8,934598	3
168	0,716866	2,389817	2,709874	2,138012	6,366159	4,533546	6,206496	3
169	0,691452	2,235717	3,850305	2,310113	-7,67578	5,293957	-7,39241	3
170	0,654779	1,861483	2,744911	2,267745	-6,55427	5,028833	6,1051	3
171	0,643755	1,895818	2,840772	2,516988	-6,81748	-5,19481	-6,76134	3
172	0,733546	2,873721	3,457077	2,151373	7,280127	5,300625	-7,92895	3
173	0,683093	2,105601	3,001679	2,464984	7,987782	6,013985	-7,56095	3
174	0,683996	1,896785	3,035479	2,428359	6,906994	4,839327	7,303747	3
175	0,582464	1,717884	2,063608	2,121987	6,110416	4,468064	6,172902	3
176	0,625552	1,842653	3,681501	1,832642	8,437115	5,149287	7,986307	3
177	0,685713	1,889697	3,414011	2,179747	7,357836	4,48896	7,296425	3
178	0,68945	2,781532	3,077072	1,793651	-7,18755	5,208864	7,364441	3
179	0,740824	2,347416	3,640174	2,432123	8,91505	6,159826	-9,52846	3
180	0,716246	2,27766	3,170697	1,741375	7,658131	-5,6244	-7,99079	3
181	0,728006	2,246661	3,288737	2,197159	-8,16949	-5,61055	7,36829	3
182	0,749803	2,723662	3,017996	2,376819	8,167194	-6,26134	-7,32727	3
183	0,751142	2,724913	2,941601	2,48417	-7,69879	5,996802	-6,73716	3
184	0,748049	2,68022	3,019222	2,384405	7,887012	-5,859	-7,26491	3
185	0,747521	2,665066	2,992719	2,416842	8,466475	-5,69873	-7,0668	3
186	0,748546	2,682104	3,014012	2,403346	8,129319	-5,67848	-7,14381	3
187	0,746849	2,623519	2,973337	2,448412	-8,36321	-5,92313	-7,04474	3
188	0,747989	2,622056	2,989097	2,459932	-8,15599	-6,10817	-7,0347	3
189	0,74683	2,627661	2,975468	2,450156	-8,28095	-6,09377	-7,05277	3
190	0,752122	2,783399	2,916345	2,101109	7,176463	5,245268	7,145658	3

241	0,695978	2,340761	3,080743	1,724886	-6,82962	-4,45086	6,705958	4
242	0,727942	2,318397	4,918513	2,277412	-7,61543	5,237589	8,179846	4
243	0,579932	1,638497	1,861674	1,470783	4,086778	3,025332	4,84045	4
244	0,662521	2,463873	2,533016	1,854525	4,835485	3,784765	5,286136	4
245	0,73867	2,289895	3,458007	1,883805	-9,16063	-5,99403	8,388581	4
246	0,54938	1,6279	1,000104	0,922116	2,18914	1,772736	-2,23389	4
247	0,720729	2,359794	3,028561	1,996109	5,340016	4,063224	-6,42798	4
248	0,758651	2,588719	3,651568	2,42916	-7,31634	-4,94685	7,810748	4
249	0,729374	2,300973	3,669664	2,426905	7,974476	-6,16668	-8,68059	4
250	0,70472	2,218451	3,389322	1,698985	7,405182	4,670165	-7,58901	4
251	0,611296	2,196238	1,996968	1,303074	3,925429	3,270872	-5,07231	4
252	0,669182	2,237735	3,708884	1,619766	-5,7464	3,791716	5,88339	4
253	0,398849	2,496034	0,343548	0,648593	-1,36595	-2,29313	1,9228	4
254	0,731567	2,29775	3,711306	1,974349	9,153655	-4,99824	7,59686	4
255	0,768231	3,228406	4,266255	2,127493	-7,87611	-5,59535	8,159004	4
256	0,708328	2,356165	3,675051	1,769644	8,068225	5,145048	7,610287	4
257	0,295679	1,054605	0,552131	0,176137	1,020925	0,927201	0,965402	4
258	0,697581	2,27989	3,180731	1,591188	-5,835	4,08794	5,682403	4
259	0,737498	2,513597	3,171941	1,839915	6,176974	4,553481	-6,81871	4
260	0,640974	2,423452	2,996039	1,899501	5,534789	3,960963	5,790246	4
261	0,753021	2,824447	3,360249	1,906942	-8,33362	-5,96668	-8,08094	4
262	0,75532	2,426726	4,64081	1,961856	-8,34009	5,31083	8,619805	4
263	0,299977	0,874226	-0,14136	-0,03368	0,016545	0,332642	0,111183	4
264	0,752606	2,515909	4,340985	1,901236	-6,89819	4,513239	6,983647	4
265	0,557746	1,541842	2,060693	1,779784	4,262325	3,218127	4,767006	4
266	0,568939	1,867636	1,457477	1,102015	2,776692	2,353336	4,677051	4
267	0,744727	2,790177	3,275674	2,309606	-6,77196	-4,82439	7,351393	4
268	0,559631	1,663453	1,135613	0,963936	2,357929	2,048247	-3,73057	4
269	0,733879	2,456674	3,838325	2,105162	7,580938	-7,01655	-8,01886	4
270	0,218949	0,83895	-0,23329	0,136668	-0,11828	0,247617	0,183787	4

321	0,188638	0,412447	1,411361	1,437799	2,561093	1,749008	4,094471	5
322	0,268084	0,595732	1,431001	1,554724	2,945253	2,011253	3,994149	5
323	0,224887	0,496852	1,615577	1,798711	3,509125	2,319767	4,071352	5
324	0,254475	0,561392	1,347446	1,420761	2,576392	1,816081	3,873596	5
325	0,191378	0,440762	1,075553	1,295281	2,455769	1,768982	2,989223	5
326	0,213588	0,489935	1,031669	1,183282	2,146182	1,597737	2,987491	5
327	-0,06257	0,530805	0,35457	-0,11563	0,979689	0,309993	0,314806	5
328	0,195368	0,428344	1,515215	1,563005	2,87458	1,929734	4,065837	5
329	0,181474	0,417395	1,054054	0,893125	2,203323	1,711728	3,131958	5
330	0,199006	0,437373	1,608683	1,541218	3,270128	2,22015	4,097993	5
331	0,185506	0,414441	1,428441	1,194444	3,392032	2,589353	3,641513	5
332	0,177244	0,39638	1,32411	1,081192	2,432307	1,517302	-3,45677	5
333	0,182202	0,408351	1,204565	1,37948	2,590871	1,716632	3,820397	5
334	0,187044	0,416424	1,282687	1,451062	2,726466	1,796889	3,903906	5
335	0,186252	0,421436	1,235273	1,277459	2,70165	2,001384	3,164602	5
336	0,20364	0,445161	1,268625	1,279367	2,24514	1,571005	5,147534	5
337	0,18822	0,415505	1,377555	1,531224	2,84475	1,877636	3,943141	5
338	0,194397	0,427181	1,509407	1,632199	3,096204	1,987242	4,422699	5
339	0,164341	0,445264	2,235774	0,450344	-3,51039	1,830577	-3,49542	5
340	0,165051	0,364214	1,409277	0,401781	2,9447	1,669504	-4,53906	5
341	0,179546	0,492269	2,3425	0,435312	-3,38155	1,648209	-3,26551	5
342	0,172637	0,386046	1,481885	0,357823	3,163884	1,817729	4,286964	5
343	0,153443	0,410972	0,037034	0,427168	-3,65652	2,163906	-4,68429	5
344	0,18473	0,41193	1,805132	0,42772	3,978572	2,276183	5,082588	5
345	-0,02356	0,139966	0,182189	0,165009	0,71505	0,5478	0,51417	5
346	0,0896	0,202963	1,96747	0,189521	4,221155	2,260585	5,706911	5
347	0,133986	0,296905	1,762974	0,27082	3,738811	2,063428	4,805083	5
348	0,10487	0,234569	1,616335	0,228465	3,369216	1,827872	5,615801	5
349	0,034421	0,08744	1,28041	0,220234	2,682777	1,3801	3,370238	5
350	0,038054	0,09369	1,97368	0,295036	4,302088	2,207098	4,530495	5

4. Phương pháp so khớp mẫu:

4.1. Đối sánh mẫu:

- Được thực hiện bằng cách so sánh các vector đặc trưng thử nghiệm với các vector đặc trưng đào tạo để xác định sự tương đồng
- Sự giống nhau giữa hai điểm dữ liệu được đo bằng khoảng cách

For a point (x_1, x_2, \dots, x_n) and a point (y_1, y_2, \dots, y_n) ,

$$\text{Manhattan distance} = 1\text{-norm distance} = \sum_{i=1}^n |x_i - y_i|$$

$$\text{Euclidean distance} = 2\text{-norm distance} = \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{1/2}$$

$$d = \sqrt{(S11 - Sn1)^2 + (S12 - Sn2)^2 + \dots + (S17 - Sn7)^2}$$

4.2. Cách thực hiện:

- Từ đặc trưng Hu ta thu được ta trích ra 80 ảnh là Train và số còn lại là để Test như vậy ta thu được 400 giá trị Hu của các quả để Train và số còn lại để Test, với quy ước:

1: Tỏi	4: Hành
2: ớt	5: Sả
3: Gừng	
- So khớp 7 giá trị vector Hu của loại thảo mộc và gia vị cần test với các vector có trong cơ sở dữ liệu
- Tính giá trị khoảng cách nhỏ nhất
- Kết luận

4.3. Kết quả:

4.3.1. Tỏi:

$$B = \begin{bmatrix} 15 & 0 & 1 & 4 & 0 \end{bmatrix}$$

4.3.1. Ớt:

$$B = \begin{bmatrix} 0 & 20 & 0 & 0 & 0 \end{bmatrix}$$

4.3.3. Gừng:

B =

1 0 29 2 0

4.3.4. Hành:

B =

1 0 2 27 0

4.3.5. Sả:

B =

2 4 2 0 12

Bảng kết quả so khớp mẫu

	Tỏi	Ớt	Gừng	Hành	Sả
Tỏi	15	0	1	1	2
Ớt	0	20	0	0	4
Gừng	1	0	29	2	2
Hành	4	0	2	27	0
Sả	0	0	0	0	12
Tổng dữ liệu test	20	20	32	30	20
Phần trăm	75%	100%	91%	90%	60%

=> Tỷ lệ nhận dạng trung bình đạt được 83%

5. Phương pháp KNN:

5.1. Tổng quan:

- KNN là thuật toán đi tìm đầu ra của một điểm dữ liệu mới bằng cách chỉ dựa trên thông tin của K điểm dữ liệu trong tập train gần nó nhất
- Thuật toán dựa trên khoảng cách giữa các đặc trưng của dữ liệu test và dữ liệu Data sau đó lấy K điểm có khoảng cách gần nhất với nó
- Công thức:

$$\text{Manhattan distance} = \text{1-norm distance} = \sum_{i=1}^n |x_i - y_i|$$

5.2. Cách thực hiện:

- Sử dụng tập train và các tập test giống như phương pháp so khớp mẫu
- Chọn K là một số tự nhiên khác 0 thích hợp
- Tìm khoảng cách của các đối tượng cần tìm label với tất cả các đối tượng trong tập train để tìm ra K đối tượng gần nhất
- Lấy label của K đối tượng gần nhất, tìm số label lớn nhất thì label đó chính là label của đối tượng test

5.3. Kết quả:

(Trong bài này, chúng em chọn K=5 để thực hiện và trình bày)

5.3.1. Tôi:

A =

12	0	2	5	1
----	---	---	---	---

5.3.2. Ốt:

A =

0	18	0	0	2
---	----	---	---	---

5.3.3. Gìrng:

A =

0	0	30	2	0
---	---	----	---	---

5.3.4. Hành: $A =$

1	0	2	27	0
---	---	---	----	---

5.3.5. Sả: $A =$

0	5	2	0	13
---	---	---	---	----

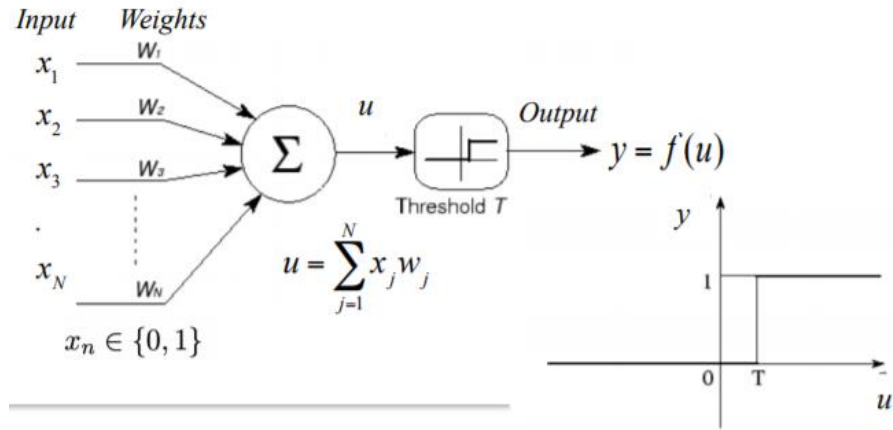
Bảng kết quả phương pháp KNN

	Tỏi	Ớt	Gừng	Hành	Sả
Tỏi	12	0	0	1	0
Ớt	0	18	0	0	5
Gừng	2	0	30	2	2
Hành	5	0	2	27	0
Sả	1	2	0	0	13
Tổng dữ liệu test	20	20	32	30	20
Phần trăm	60%	90%	94%	90%	65%

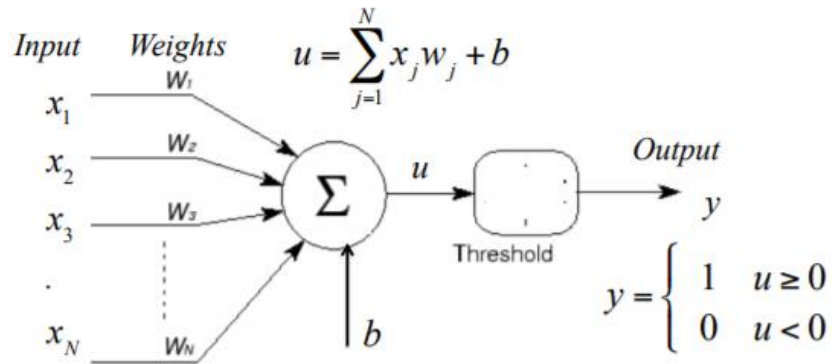
=> Tỷ lệ nhận dạng trung bình đạt được 80%

6. Phương pháp huấn luyện mạng Neural:

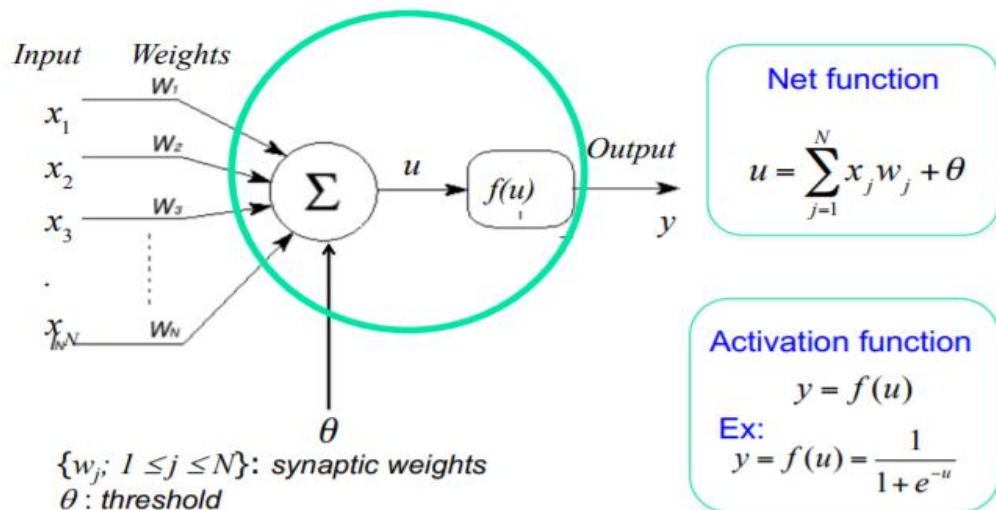
- Một mô hình tính toán toán học mô hình hóa tế bào thần kinh của con người



- Mô hình neural cảm thụ:

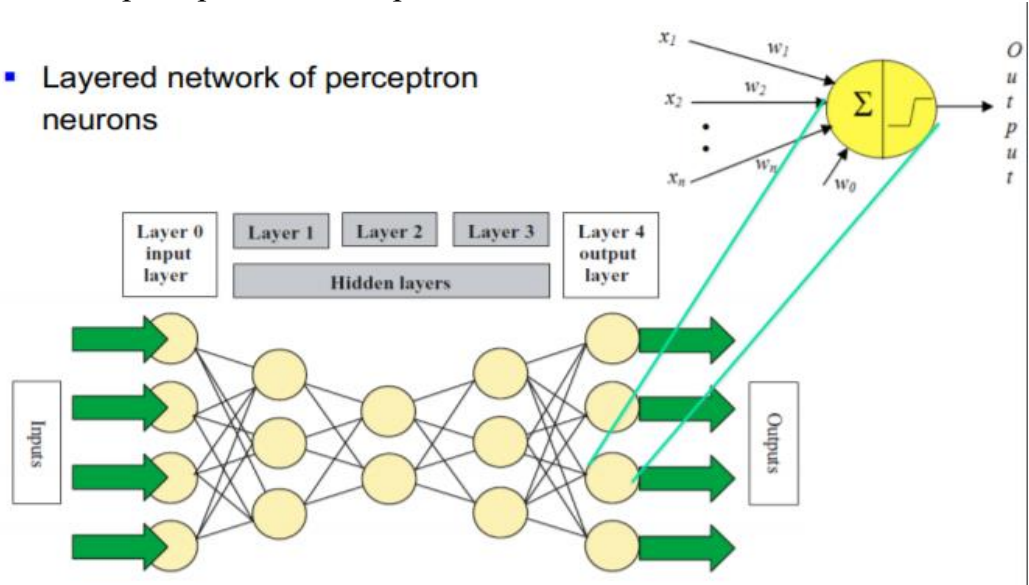


- Mô hình Neural tổng quát:



- Mô hình perceptron nhiều lớp

■ Layered network of perceptron neurons

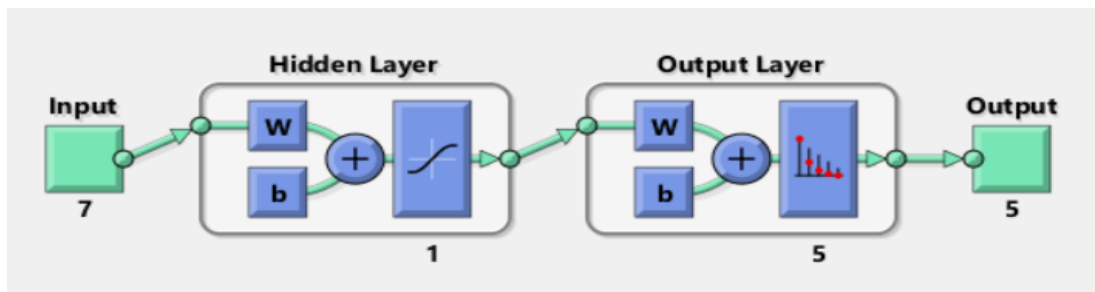


6.1. Thiết kế mạng nơron nhân tạo:

- Thu thập và biểu diễn dữ liệu
- Thiết lập cấu trúc liên kết mạng
- Tạo tham số mạng
- Khởi tạo giá trị trọng số và thiên vị

Training:	80%	418 samples
Validation:	10% ▾	52 samples
Testing:	10% ▾	52 samples

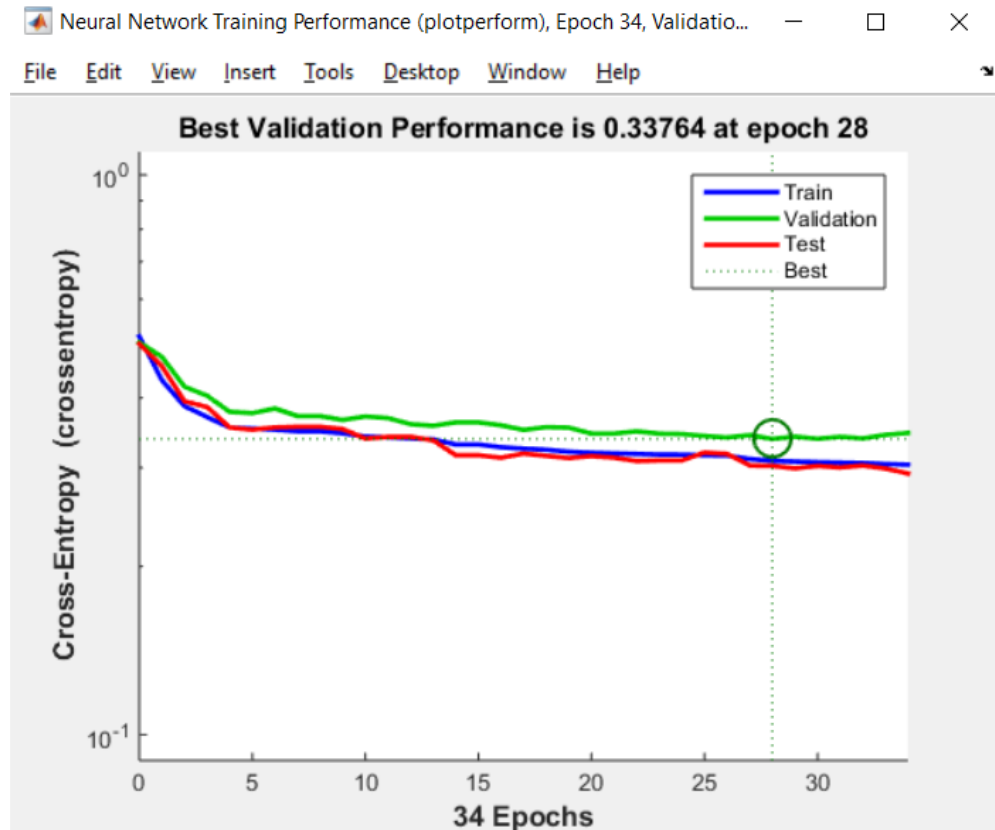
- Training
- Xác thực → thiết kế lại hoặc sử dụng



- 1 lớp ẩn
- Lớp đầu vào 7 noron (vector đặc trưng Hu)
- Lớp đầu ra 5 noron (5 class)

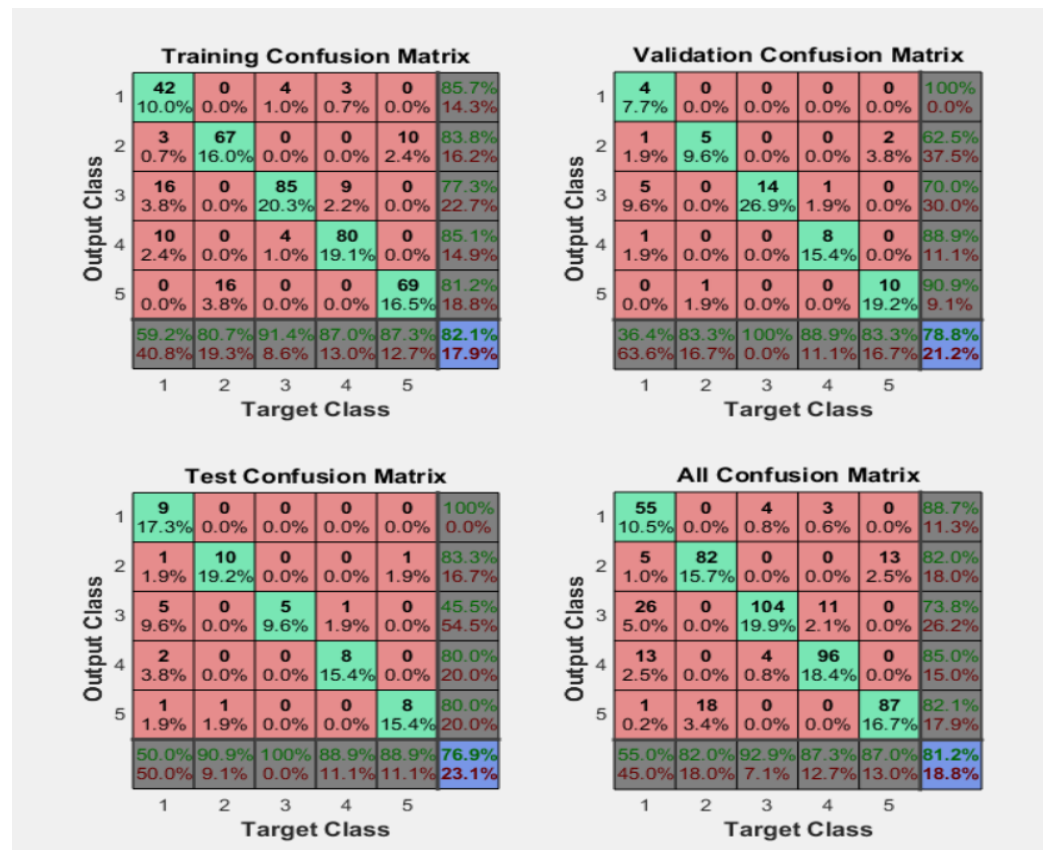
6.2. Kết quả huấn luyện:

6.2.1. Hiệu suất:



- Nhận xét:
 - Cross - entropy có xu hướng giảm khi số epochs tăng, điều này đồng nghĩa với việc mô hình được huấn luyện có hiệu quả nhận dạng dữ liệu ngày càng tốt
 - Cả 3 đường train, test, validation đều có xu hướng giảm, tức là mô hình huấn luyện không bị overfitting
 - Hiệu suất tốt nhất nằm ở epoch 28

6.2.2. Ma trận nhầm lẫn:



- Nhận xét:
 - Kết quả của lần training này đối với tập test cho thấy tỉ lệ nhận dạng đúng 77%
 - Tỉ lệ của hệ thống là 81%

II. Kết luận và hướng phát triển:

1. Sản phẩm và kết luận:

❖ Sản phẩm

- Sau khi kết thúc học phần sản phẩm thu được là:
 - Cơ sở dữ liệu
 - Code trong quá trình thực hiện
 - Báo cáo bằng slide
 - Báo cáo bằng pdf
 - Demo được trên Matlab bằng 3 phương pháp so khớp mẫu, KNN và Neural với một bức ảnh màu hoặc một tập giá trị Hu (file excel)

❖ **Kết luận:**

- Các phương pháp nhận dạng cho ra kết quả tương đối tốt đều trên 80%
- Hoàn thành các yêu cầu đề tài đặt ra, tuy nhiên số lượng ảnh thu thập còn hạn chế, một số ảnh lấy từ internet
- Đánh giá các phương pháp:
 - Phương pháp so khớp mẫu:
 - Ưu điểm: Đơn giản, ít công thức, tỉ lệ nhận dạng trung bình 82%
 - Nhược điểm: Kết quả nhận dạng từng loại chưa cao, còn bị nhầm lẫn giữa các loại có hình dạng tương tự giống nhau
 - Phương pháp KNN:
 - Ưu điểm: Đơn giản, tỉ lệ nhận dạng trung bình 80%
 - Nhược điểm: Phụ thuộc vào hệ số K, khi thử với hệ số K nhỏ thì kết quả nhạy cảm với nhiễu và K lớn thì kết quả có nhiều điểm của các loại khác có tương tự hình dáng
 - Phương pháp mạng Neural:
 - Dùng tool có sẵn trên matlab (NPRTool)
 - Kết quả cho thấy tỉ lệ nhận dạng đúng 81% thời gian Train và Test khá nhanh

2. Hướng phát triển:

- Tiếp tục phát triển đề tài, mở rộng phạm vi nghiên cứu rộng hơn
- Tăng độ chính xác của dữ liệu
- Hoàn thiện dự án phát triển thành App ứng dụng thực tế
- Phát triển nhận dạng mẫu rộng hơn về các lĩnh vực trong đời sống ngày nay và cả sau này
- Đưa machine learning tiếp cận đời sống

MỤC LỤC

Nội dung	Trang
I. Cơ sở lý thuyết và phương pháp thực hiện	4
1. Thu thập dữ liệu.....	4
2. Tiền xử lý.....	5
3. Trích đặc trưng	7
4. Phương pháp so khớp mẫu	11
5. Phương pháp KNN	14
6. Phương pháp huấn luyện mạng Neural	16
II. Kết luận và hướng phát triển	19
1. Sản phẩm và kết luận	19
2. Hướng phát triển	20

PHỤ LỤC

1. Mọi CSDL, code, các tập train, test đều có ở:
https://drive.google.com/drive/u/1/folders/1tTO1b7HULhvj2l0ef_GOdJI_RMl5sCt
2. Một số hình ảnh được sử dụng làm dữ liệu trong đề tài:
<https://github.com/Horea94/Fruit-Images-Dataset>
3. Tài liệu tham khảo:
 - Slide bài giảng bộ môn KT Nhận Dạng
 - Bài viết: <https://moingaymotchut.wordpress.com/2017/10/16/cai-dat-thuat-toan-knn-bang-matlab/>