

## ZASTOSOWANIE METODY WEKTORÓW NOŚNYCH ORAZ KOMPUTEROWEJ ANALIZY OBRAZU W KLASYFIKACJI KORZENI MARCHWI

Monika Janaszek, Jędrzej Trajer

*Katedra Podstaw Inżynierii, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie*

**Streszczenie.** W pracy poruszono zagadnienie podejmowania decyzji o przydatności przetwórczej marchwi na podstawie uproszczonej informacji o barwie jej korzeni. Sprawdzono w jakim stopniu barwa pozwoli na odwzorowanie skupień korzeni o podobnych cechach chemicznych, decydujących o dalszym przeznaczeniu surowca. Do klasyfikacji korzeni wykorzystano metodę wektorów nośnych (SVM). Barwę marchwi odczytano z cyfrowych obrazów jej korzeni. Trafność klasyfikacji w zbiorze testowym wskazuje, że barwę można wykorzystać do opracowania wielokryterialnej klasyfikacji marchwi pod względem jej przydatności przetwórczej.

**Słowa kluczowe:** analiza obrazu, klasyfikacja, marchew, SVM

### Wstęp

Marchew jest już od lat przedmiotem badań, mających na celu wskazanie cech istotnie wpływających na stopień jej akceptowalności przez konsumentów oraz źródeł zmienności tych cech. Badania te skupiają się przede wszystkim na poszukiwaniu zależności między warunkami pogodowymi w okresie wegetacji, cechami odmianowymi, warunkami uprawy, zbioru i przechowywania a jakością sensoryczną surowca. Wyniki prowadzonych badań wskazują, iż istnieją zależności między cechami fizyko-chemicznymi marchwi a jej jakością sensoryczną [Talcot i in. 2001; Varming i in. 2004]. Stwierdzono również, że cechy fizyko-chemiczne korzeni mogą zostać wykorzystane do identyfikacji odmianowej [Rosenfeld i in. 1997, Skrede i in. 1997]. Stwierdzono, że na podstawie kształtu korzeni oraz parametrów barwy, uzyskanych z obrazów korzeni marchwi możliwa jest identyfikacja odmiany marchwi. Opracowany model wykazywał jednak dużą wrażliwość na takie źródła zmienności, jak lata i miejsce uprawy [Horgan i in. 2001].

W pracy poruszono zagadnienie podejmowania decyzji o przydatności przetwórczej marchwi na podstawie informacji, pochodzącej z cyfrowych obrazów jej korzeni. Stosowane dotychczas metody oceny marchwi skupiają się głównie na analizie składu chemicznego, który jest jednym z najważniejszych determinantów, jeśli chodzi o przydatność przetwórczą tego surowca. Stosowanie analitycznych metod oceny przetwórczej przydatności marchwi jest jednak kosztowne i czasochłonne, ponadto nie każde przedsiębiorstwo w sektorze przetwórstwa owocowo-warzywnego może pozwolić sobie na utrzymanie zaplecza laboratoryjnego. W praktyce ocena taka jest często przeprowadzana w warunkach

niepełnej i niepewnej wiedzy i ogranicza się jedynie do inspekcji pod kątem zdrowotności oraz zawartości substancji szkodliwych i niepożądanych. Informacja o odmianie uznawana jest natomiast za wystarczający determinant wartości odżywczej marchwi. Stwarza to trudności i wpływa na precyzję i jednoznaczność tej oceny. Nie bez znaczenia pozostają też właściwości reologiczne korzeni, na podstawie których określa się teksturę, mającą wpływ na parametry procesów przetwórczych oraz barwa korzeni, która może ulegać zmianie w zależności od stosowanych w przemyśle przetwórczym technologii. W przypadku marchwi w stanie surowym pomiar takich cech, jak barwa czy tekstura sprawia trudności z uwagi na specyfikę budowy korzeni. Marchew ze względu na anizotropową budowę korzenia i jego morfologiczną niejednorodność stanowi trudny obiekt badawczy, jeśli chodzi o pomiar barwy. Trudno zatem określić, w którym miejscu korzenia i w ilu powtórzeniach należy dokonać pomiaru oraz jaka forma materiału poddawanego analizie barwy byłaby najbardziej odpowiednia, aby uzyskać miarodajne wyniki [Zapotoczny i Zielińska 2005]. Alternatywę mogą stanowić metody komputerowej analizy obrazu (ang. *Digital Image Analysis – DIA*), umożliwiające pozyskanie parametrów korzeni w sposób obiektywny, powtarzalny i wystandaryzowany [Abbott 1999]. Ze względu na większą dokładność pomiaru, porównywalność ocenianych parametrów oraz skrócenie czasu trwania analiz, komputerową analizę obrazu uznaje się za alternatywę dla stosowanych dotąd konwencjonalnych metod oceny produktów biologicznych [Du i Sun 2004].

## Cel pracy

Celem pracy było sprawdzenie czy na podstawie uproszczonej informacji o barwie, można odwzorować skupienia obiektów o podobnych cechach chemicznych, które decydują o dalszym przeznaczeniu surowca.

## Materiał i metody

Materiałem badawczym były korzenie marchwi jadalnej (*Daucus carota* L.) odmian: Florida, Kathmandu, Kazan, Laguna, Mazurska, Recoleta, Sugarsnax, pochodzące z plantacji RZD w Żelaznej, z dwóch kolejnych lat zbiorów, 2005 i 2006.

W każdym roku badań wykonano analizę składu chemicznego korzeni oznaczając: masę suchej substancji (SM), popiół (POP), zawartość pektyn (PEK), ekstrakt ogólny (EKS), cukry ogółem (C\_ OG) i redukujące (C\_RED), witaminę C (WIT\_C), karotenoidy ogółem (KAR\_ OG) i  $\beta$ -karoten (B\_KAR).

Parametry barwy korzeni marchwi otrzymano z obrazów ich przekrojów wzdłużnych, utrwalonych aparatem cyfrowym OLYMPUS 5050Z. Korzenie umieszczano na powierzchni, wykonanej z białego, przepuszczającego i równomiernie rozpraszającego światło poliwęglanu. Powierzchnia fotografowanego obiektu oświetlana była źródłem światła o temperaturze barwowej 5000 K. Przed wykonaniem zdjęć dokonywano kalibracji aparatu, poprzez ustawienie balansu bieli. Wszystkie zdjęcia wykonano w kolorze, przy czułości 100 ISO, ogniskowej 21,3 mm (odpowiednik ogniskowej 105 mm w aparacie małoobrazkowym 35 mm) oraz przysłonie f2,8 i zapisano w formacie ORF (ang. Olympus RAW

File). Zdjęcia zostały wczytane z karty pamięci aparatu do komputera, za pomocą programu Adobe® Camera Raw [Adobe Systems Inc. 2003]. Dalsze przetwarzanie obrazów odbywało się w trybie wsadowym. Z obrazów usunięto tło, co było równoznaczne z wyodrębnieniem z tła widocznych na obrazach korzeni marchwi. Zdjęcia zostały zapisane jako mapy bitowe o rozmiarach 2560×1920 pikseli i rozdzielczości 300 dpi, a następnie przetworzone, przy pomocy programu autorskiego, napisanego w języku C#, do postaci tekstowej w celu wczytania zawartych w nich danych do oprogramowania analitycznego. Przetwarzanie polegało na odczytaniu jasności pikseli w czerwonej, zielonej i niebieskiej warstwie mapy bitowej i zapisaniu tych informacji do pliku. Jasność pikseli w kolejnych rzędach mapy bitowej była odczytywana począwszy od jej górnego lewego rogu. W wyniku takiego postępowania, dla map bitowych otrzymano pliki źródłowe, które następnie zostały wczytane do programu SAS<sup>TM</sup> [SAS Institute Inc. 2004]. Tabele, utworzone przez procedury importu w bibliotece programu SAS<sup>TM</sup> połączono w jedną tabelę, zawierającą w kolejnych polach informacje o roku zbioru, odmianie, numerze obiektu oraz wartościach składowych R, G i B.

W celu pominięcia w analizach obszarów powstałych po wycięciu tła z map bitowych, z tabeli danych usunięto rekordy, w których wartości składowych R, G i B jednocześnie wynosiły 255 (piksele o barwie białej). Pozostałe w zbiorze wartości zmiennych R, G, B, odpowiadające pojedynczym pikselom obrazu korzenia, wykorzystano do wyznaczenia wektorów średnich R, G, B dla każdego korzenia. Zbiór danych po tym etapie zawierał 140 obserwacji.

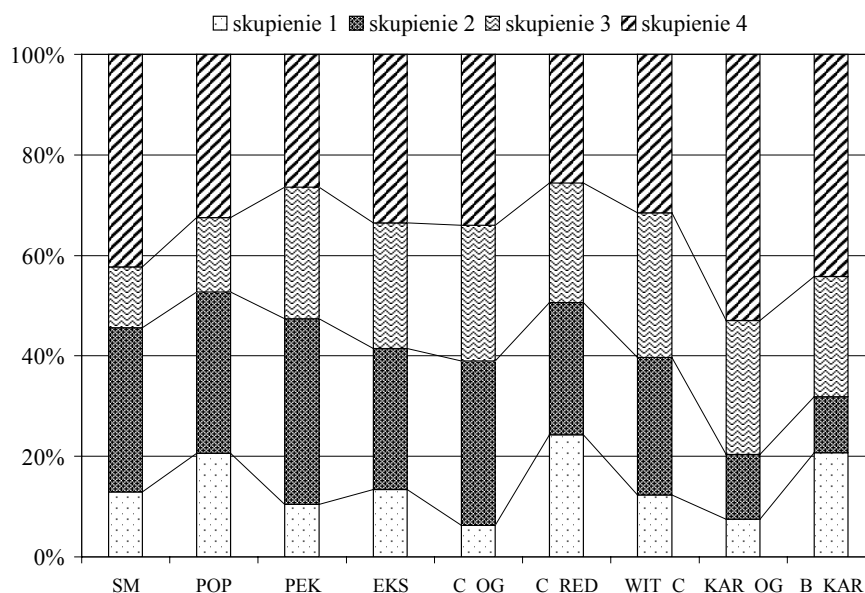
Ze względu na brak jasno sprecyzowanych i sformalizowanych wymagań co do cech surowca przeznaczanego na określone cele przetwórcze, klasyfikację poprzedzono wyodrębnieniem grup obiektów o podobnym składzie chemicznym. Pojedyncza grupa (skupienie) stanowiła klasę obiektów, którą wykorzystano jako wzorzec w dalszych analizach. Jako narzędzie klasyfikacyjne wykorzystano metodę wektorów nośnych typu drugiego ( $\nu$ -SVM), z radialnym typem jądra [Vapnik 1995], która realizowana była przez sieć neuronową. Dane wprowadzane do sieci podzielono na dwa zbiory: uczący i testowy. Zbiory te zawierały odpowiednio 75% i 25% obserwacji. Na wejścia sieci wprowadzono ciąg uczący, w którego skład weszły wektory średnich R, G, B. Uczenie modelu SVM poprzedzono dziesięciokrotną walidacją krzyżową, w celu doboru optymalnych wartości parametrów sieci  $\nu$  oraz  $\gamma$ . Proces uczenia sieci obejmował tysiąc iteracji, a warunkiem zatrzymania uczenia było osiągnięcie błędu na poziomie 0,001.

## Omówienie wyników

Charakterystykę wzorców dla sieci SVM, wyłonionych na podstawie składu chemicznego korzeni, prezentuje rysunek 2.

W nawiązaniu do wyników badań, szczególnie nad przydatnością przetwórczą marchwi [Talcott i in. 2001; Borowska i in. 2005; Zielińska i in. 2005], scharakteryzowano wstępnie wyłonione wzorce, poprzez obserwację w ich obrębie średnich zawartości oznaczanych substancji. Skupienie pierwsze charakteryzowało się niskimi zawartościami oznaczonych substancji, co raczej nie czyni zawartej w nim marchwi przydatną do przetwórstwa, głów-

nie ze względu na niską masę suchej substancji oraz najmniejszą zawartość cukrów ogółem i karotenoidów. Ze względu na wysoką (lecz porównywalną do pozostałych wzorców) zawartość cukrów redukujących nie jest to dobry surowiec na susz. Surowiec tworzący skupienie drugie odpowiadał marchwi przydatnej do produkcji soków przecierowych. Pożądane cechy w tym przypadku to wysoki ekstrakt, masa suchej substancji i zawartość cukrów ogółem. Surowiec w skupieniu trzecim, mimo niskiej masy suchej substancji, mógłby być przydatny do produkcji mrożonek. Do cech pożądanych w tym przypadku należą wysoki ekstrakt, zawartość cukrów ogółem, pektyn i karotenoidów, a korzenie w tym skupieniu charakteryzowały się przeciętną zawartością tych substancji. Marchew zaliczona do skupienia czwartego posiadała parametry pożądane w produkcji soków zagęszczonych, gdzie surowiec powinien charakteryzować się wysokim ekstraktem, zawartością cukrów i karotenoidów.



*Źródło: obliczenia własne*

Rys. 1. Charakterystyka skupień (wzorców SVM) pod względem składu chemicznego  
 Fig. 1. Characteristics of clusters (SVM patterns) in terms of chemical composition

Zadaniem sieci SVM było prawidłowe rozpoznanie wzorców marchwi, wyłonionych ze względu na skład chemiczny, jedynie na podstawie danych o barwie przekroju korzeni. Po dziesięciokrotnej walidacji krzyżowej dobrano dla sieci SVM parametry  $\nu$  oraz  $\gamma$ , których wartości wyniosły odpowiednio 0,4 i 0,1. Jakość działania sieci oceniono na podstawie

błędów, z jakimi rozpoznała ona przynależność obiektów ze zbioru testowego do zdefiniowanych wcześniej wzorców, czyli na podstawie tzw. macierzy pomyłek. Procentowe udziały przypadków testowych poprawnie przypisanych do wzorców były następujące: skupienie 1 – 96,67%, skupienie 2 – 88,89%, skupienie 3 – 90%, skupienie 4 – 100%. Wysoka trafność klasyfikacji w przypadku skupienia 4 wynikała z małej liczności tego skupienia. Tego typu sytuacji można w przyszłości zapobiegać przez stosowanie tzw. próbkowania w procesie wylaniania zbioru testowego. W tym wypadku do wskazania zbioru testowego wykorzystano standardowy generator liczb pseudolosowych. Ogólna trafność klasyfikacji w zbiorze uczącym wyniosła 91,43%, a w zbiorze testowym 94,29%. Liczba wektorów związanych wyniosła 7, co oznacza, że algorytm uczący dokonał poprawnej separacji wzorców.

## Podsumowanie

Wyniki klasyfikacji wskazują, że informacja o barwie wprowadzona do sieci w postaci wektora średnich była informacją wystarczającą do tego, aby prawidłowo klasyfikować wyłonione wcześniej wzorce. Uzyskana trafność klasyfikacji, przekonuje, iż na podstawie barwy można odwzorować korzenie marchwi, charakteryzujące się podobnymi cechami składu chemicznego. Analiza składu chemicznego oraz barwy może być zatem wykorzystana do opracowania wielokryterialnej klasyfikacji marchwi pod względem jej przydatności przetwórczej. Należy również przypuszczać, że tego typu klasyfikacja nie będzie oparta na zależnościach między pojedynczymi cechami składu chemicznego i składowymi barwy. Barwa stanowi bowiem odwzorowanie ogólnej informacji o składzie chemicznym korzeni, a nie o zawartości konkretnej substancji (np. karotenoidów).

Zastosowanie sieci SVM w zagadnieniu klasyfikacji korzeni marchwi potwierdziło jej wysoką skuteczność, o czym świadczy trafność klasyfikacji w zbiorze testowym. Przydatność sieci SVM w rozwiązywaniu problemów klasyfikacji wynika z samodzielnego i optymalnego doboru struktury sieci, zapewniającego zdolności uogólniające. Ograniczona liczba wektorów podtrzymujących, a tym samym zredukowana struktura sieci, umożliwia skuteczne jej zastosowanie w mniej licznych zbiorach.

## Bibliografia

- Abbott J. A.** 1999. Quality measurement of fruits and vegetables. *Postharvest Biology and Technology* 15. s. 207-225.
- Borowska E. J., Zadernowski R., Szajdek A., Majewska K., Budrewicz G.** 2005. Cechy organoleptyczne, fizyczne i chemiczne wybranych odmian marchwi przydatnych do produkcji soku. *Polish Journal of Natural Sciences*, 18(1). s. 174-186.
- Du C.-J., Sun D.-W.** 2004. Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in Food Science & Technology* 15. s. 230-249.
- Horgan G. W., Talbot M., Davey J. C.** 2001. Use of statistical image analysis to discriminate carrot cultivars. *Computers and Electronics in Agriculture* 31. s. 191-199.

- Rosenfeld H. J., Baardseth P., Skrede G.** 1997. Evaluation of carrot varieties for production of deep-fried carrot chips-IV. The influence of growing environment on carrot raw material. *Food Research International* 30(8). s. 611-618.
- Skrede G., Nilsson A., Baardseth P., Rosenfeld H. J., Enersen G., Slinde E.** 1997. Evaluation of carrot varieties for production of deep-fried carrot chips-III. Carotenoids. *Food Research International* 30(1). s. 73-81.
- Talcott S. T., Howard L. R., Brenes C. H.** 2001. Factors contributing to taste and quality of commercially processed strained carrots. *Food Research International* 34. s. 31-38.
- Vapnik V.N.** 1995. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, Inc. ISBN 0-387-98780-0.
- Varming C., Jensen K., Møller S., Brockhoff P. B., Christiansen T., Edelenbos M., Bjørn G. K., Poll L.** 2004. Eating quality of raw carrots-correlations between flavour compounds, sensory profiling analysis and consumer liking test. *Food Quality and Preference* 15. s. 531-540.
- Zapotoczny P., Zielińska M.** 2005. Rozważania nad metodyką instrumentalnego pomiaru barwy marchwi. *Żywność* 1(42). s. 121-132.
- Zielińska M., Zapotoczny P., Markowski M.** 2005. Colour standard and homogenous groups of dried carrots of 34 commercial varieties. *Polish Journal of Food and Nutrition Sciences*, 14/55 (1). s. 51-56.
- Adobe Systems Incorporated. 2003. Adobe Photoshop CS. California 95110, 345 Park Avenue, San Jose, USA.
- SAS Institute Inc. 2004. SAS 9.1 Companion for Windows. Cary, NC, USA: SAS Publishing, SAS Institute Inc.

## APPLICATION OF SUPPORT VECTOR MACHINES AND DIGITAL IMAGE ANALYSIS IN CARROT ROOTS CLASSIFICATION

**Abstract.** The article presents the study concerning the question of deciding on the processing suitability of carrot on the basis of simplified information about the color of roots. A possibility of mapping clusters of carrot roots having a similar chemical composition, which determine the further allocation of raw material, was examined. In classification of the roots support vector machine (SVM) was used. Carrot color was read from a digital image of its roots. Classification accuracy in the test set indicates that the color can be used to develop a multi-classification of carrots in terms of its processing suitability.

**Key words:** digital image analysis, classification, carrot, SVM

**Adres do korespondencji:**

Monika Janaszek; e-mail: monika\_janaszek@sggw.pl  
Katedra Podstaw Inżynierii  
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie  
ul. Nowoursynowska 166  
02-787 Warszawa