

ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ АСК-АНАЛИЗА И ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ЭЙДОС ДЛЯ ИДЕНТИФИКАЦИИ СОРТОВ И ГИБРИДОВ ХРИЗАНТЕМЫ ПО КОНТУРУ ЛИСТА

Якушина Л.Г., Киселёва Н.С., Слепченко Н.А.

*Федеральное государственное бюджетное учреждение науки
«Федеральный исследовательский центр «Субтропический научный центр
Российской академии наук»,
г. Сочи, Россия, e-mail: vishnya584@yandex.ru*

Важной задачей селекции хризантемы является идентификация различных сортов и гибридов, в том числе по контурам листьев, так как в форме контура конкретного листа содержится информация о сорте: как о генотипе, так и о фенотипе. В качестве инструментария для построения измерительной системы и её применения предлагается автоматизированный системно-когнитивный анализ (АСК-анализ) и её программный инструментарий интеллектуальная система «Эйдос» с использованием сканированных изображений листьев. В работе показаны результаты количественной оценки внешнего контура листа сортов и гибридов хризантемы, на основе которой сформированы обобщённые образы классов изучаемых генотипов по форме листа и изучено влияние значений их генотипических особенностей. С помощью наиболее достоверной модели INF4, при интегральном критерии «Сумма знаний», решены задачи количественной идентификации разных сортов и гибридов по контурам листьев. Изучаемые сортообразцы хризантемы коллекции ФИЦ СНИЦ РАН разделены на два основных кластера, включающие более мелкие группы, по детерминируемым ими признакам контура листа.

Ключевые слова: хризантема, сорт, гибрид, системная идентификация изображения, лист, генотип, сходство-различие, обобщенные образы, система «Эйдос», АСК-анализ.

Хризантема садовая (*Chrysanthemum × hortorum* Bailey) культивируется в условиях Черноморского побережья Краснодарского края более 45 лет. За эти годы в ФИЦ СНИЦ РАН прошли испытания более 40 мелкоцветных и более 150 сортов крупноцветных хризантем [16]. Хризантема используется для озеленения, на срез и как горшечное растение. Соответственно растения обладают множеством разнообразных признаков: высокими декоративными качествами, разнообразными формами куста и листьев, но, к сожалению, многие возделываемые сорта имеют неместное происхождение, а значит слабо адаптированы, и со временем «стареют», вырождаются, снижается эффективность их возделывания.

Актуальной задачей в настоящее время является создание сортов, адаптированных к условиям возделывания, жаростойких, устойчивых к болезням и вредителям [11].

В Федеральном исследовательском центре «Субтропический научный центр Российской академии наук» (далее – ФИЦ СНЦ РАН) для получения отечественных сортов хризантемы, которые не только не уступают зарубежным по декоративности и продуктивности, но и являются более устойчивыми к абиотическим факторам среды, проводятся межсортовые скрещивания [11, 16, 15]. Родительские формы должны обладать ярко выраженными хозяйственно-ценными и декоративными признаками. При выборе исходных форм критериями служат: устойчивость к условиям произрастания, оригинальность формы и окраски соцветия, продолжительность цветения, форма листа и куста, устойчивость цветоноса, высокая декоративность. Подбор родителей с необходимым генотипом является сложной задачей, так как фенотип контролируется зачастую несколькими генами и их сочетание в потомстве трудно предугадать. Поэтому необходим поиск количественной информации для осуществления селекционного процесса, а именно, определения вероятного генотипа родителей при подборе пар [1, 2].

Хризантема имеет очень сложную гетерозиготную природу, поэтому в гибридном потомстве наблюдается большое разнообразие форм по морфологическим и биологическим признакам. При вегетативном размножении наблюдаются различные изменения. Кроме этого, высокий уровень фенотипической изменчивости создаёт проблему количественного определения различия и сходства гибридных сеянцев.

При помощи разных систем можно получить количественные характеристики фенотипических признаков гибридных сеянцев, сортов, мутантов и др. [6, 8, 14]. Используя кластерный анализ проводят сравнение по контурам листьев сортообразцов различных садовых и цветочных культур. Форма листа зависит не только от генотипа, но и от множества различных внешних и внутренних факторов. Данное явление носит название «полиморфизм». Все эти факторы будут оказывать влияние на результаты анализа листьев [6, 8].

В литературе имеются примеры описания и сравнения обобщённых образов сортов, сформированных путём многопараметрической типизации [6, 7]. Подобное обобщение сводит к минимуму влияние полиморфизма и выносит на передний план генетические особенности. В частности, форма контура отдельного листа расценивается как количественная информация о сорте и содержит информацию как о генотипе (подлинная форма листа), так и о фенотипе (изменения, проявившиеся в результате воздействия окружающей среды).

Цель исследований – оценить возможность применения автоматизированного системно-когнитивного анализа (АСК-анализа) и его программного инструментария – системы «Эйдос», как интеллектуальной измерительной технологии для решения актуальных задач в селекции хризантемы садовой (*Chrysanthemum × hortorum*), в частности, количественной идентификации разных сортов и гибридов по контурам листьев (установление степени схожести или различия по комплексу морфологических признаков листьев).

Фенологические наблюдения, особенности роста и развития гибридных сеянцев проводили в опытных посадках ФИЦ ССНЦ РАН (г. Сочи) согласно методикам Госсортоиспытания сельскохозяйственных культур [10, 13, 15] и проведения испытаний на отличимость, однородность и стабильность хризантемы (многолетней) *Chrysanthemum* sp. [11].

Объектами исследований являлись сортообразцы из генетической коллекции рода хризантема садовая (*Chrysanthemum × hortorum*) (виды, сорта и гибриды межсортовых и межвидовых скрещиваний), сохраняемые в ФИЦ ССНЦ РАН. В качестве исходных данных о сортах и гибридах хризантемы для проведения исследования использовались сканированные изображения листьев. В качестве инструментария применялся АСК-анализ и система «Эйдос» (открытое программное обеспечение: http://lc.kubagro.ru/aidos/_Aidos-X.htm) для построения измерительной системы и её применения.

Листья являются наиболее удобным объектом исследования и содержат информацию о геноме и, следовательно, других фенотипических признаках, которые зависят от генотипа и условий окружающей среды. Материалом исследования служили листья средней части побегов 26 сортов и 22 гибридов хризантемы. Листья были сканированы, измерены и проанализированы методом АСК-анализа с применением его программного инструментария – интеллектуальной системы «Эйдос» [3, 6, 7].

Решение поставленной задачи выполняется через ряд этапов. Сначала сканированные изображения листьев записываются в виде графических файлов, а затем оцифровываются в режиме 2.3.2.4 по их внешним контурам.

Формирование моделей обобщённых образов листьев различных сортов и гибридов проводилось на основе конкретных листьев (многопараметрическая типизация), в результате чего были получены обобщённые образы листьев сортов и гибридов. Пример контурного изображения листа показан на рисунке 1.

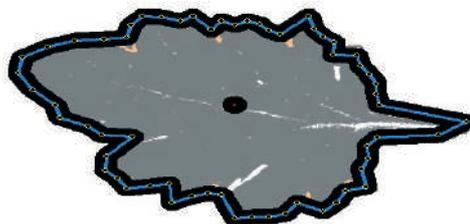


Рис. 1. Пример контурного изображения листа для сорта хризантемы садовой ‘Annecy White’

В дальнейшем, режим 2.3.2.4 системы «Эйдос» формирует Excel-таблицу, в которой содержатся результаты оцифровки изображений. Кроме того, формируется таблица с усреднёнными данными по классам. Программой определяется количество заданных текстовых и числовых классификационных и описательных шкал и градаций и, после выхода на создание модели, начинается процесс импорта данных оцифровки изображений в базы данных системы «Эйдос». При этом создаются справочники классификационных и описательных шкал и градаций, исходные данные кодируются с их использованием, в результате чего формируется обучающая выборка и база событий (эвентологическая база данных).

Полученные результаты и их обсуждение. В результате работы режима 2.3.2.4 была сформирована одна классификационная шкала с суммарным количеством градаций (классов) 59 и 60 описательных шкал с суммарным числом градаций 600. С использованием классификационных и описательных шкал и градаций исходные данные были закодированы и в результате получена обучающая выборка, которая представляет собой нормализованные исходные данные, то есть таблицу исходных данных, закодированную с помощью вышеуказанных шкал [4, 6, 7].

Синтез и верификация (оценка достоверности) 3 статистических моделей (корреляционная матрица, матрицы условных и безусловных процентных распределений) и 7 системно-когнитивных моделей (моделей знаний) были проведены на выборке из 302 листьев 59 сортов и гибридов.

Оценка достоверности моделей в системе «Эйдос» осуществляется путём решения задачи классификации объектов обучающей выборки по обобщённым образам классов и подсчёта количества истинных положительных и отрицательных, а также ложных положительных и отрицательных решений по F-мере Ван Ризбергена, а также по критериям L1- и L2-мерам проф. Е.В. Луценко, смягчающие и преодолевающие недостатки F-меры [5]. Определение наиболее

достоверной модели и придание ей статуса текущей в режиме 4.1.3.6 показало, что очень многие модели имеют максимальную возможную достоверность по критерию L2, но наиболее подходящей моделью по интегральному критерию: «Сумма знаний» можно считать INF4 (рис. 2). Поэтому исследование моделируемой предметной области путём исследования этой модели корректно можно считать исследованием самой моделируемой предметной области.

Из рисунка 2 видно, что достоверность модели INF4 с интегральным критерием «Сумма знаний» по метрике профессора Е.В. Луценко $L2 = 0,921$ (при максимуме 1,000), что является высоким показателем, причём достоверность правильной идентификации составляет 100 %. Интегральный критерий «Сумма знаний» представляет собой суммарное количество знаний, содержащееся в системе факторов различной природы, характеризующих сам объект управления, управляющие факторы и окружающую среду, о переходе объекта в будущие целевые или нежелательные состояния [3].

Частотные распределения числа истинных и ложных положительных и отрицательных решений по результатам прогнозирования идентификации по листьям сортов и гибридов хризантемы с учётом их генотипических особенностей в модели INF4 по данным обучающей выборки показаны на рисунке 3 и также подтверждают, что созданная модель не даёт ложных решений. Рисунок 3 содержит изображения двух частотных распределений, похожих на нормальные, сдвинутых относительно друг друга по фазе. Левое распределение, меньшее по амплитуде, включает истинно-отрицательные и ложно-положительные решения, а правое, большее по амплитуде, включает ложные отрицательные и истинно-положительные решения. Сдвиг этих распределений относительно друг друга и другие различия между ними и позволяют решать задачу прогнозирования и другие задачи. В нашем случае видно, что модель INF4 лучше определяет непринадлежность объекта к классу, чем принадлежность (отрицательные решения всегда истинные). В положительных решениях до 0 идут только ложные решения, при уровнях сходства от 0 до 20 % – ложные и истинные, после 20 % встречаются только истинные решения.

Степень детерминированности класса тем выше, чем больше среднее количество информации в различных значениях факторов о переходе объекта моделирования в состояние, соответствующее классу. Парето-кривая степени детерминированности классов (градации классификационных шкал) (рис. 4), наглядно показывает, что в модели INF4 около 50 % наиболее значимых классов обеспечивают 84 % суммарной значимости. В соответствии со схемой обработки данных модель INF4 определена как текущая.

4.1.3.6. Обобщ. форма по достов. моделям при разн. инт. крит. Текущая модель: "INF1"

Наименование модели и частного критерия	Интегральный критерий	С. у. н. р.	Средний модуль условной оценки истинно-отрицат. решений	С. у. р.	С. у. д.	A-Точность модели AРrecision = ATR/ATR+	A-Полнота модели AРrecall = ATR/ATR+	L2-мера проф. E.V. Луценко	Процент правильной идентификац...
1. ABS - частный критерий: количество встреч сочетаний "клас...	Корреляция абс частот с обр...	0.104				0.704	1.000	0.826	100.000
1. ABS - частный критерий: количество встреч сочетаний "клас...	Сумма абс частот по признак...					0.722	1.000	0.839	100.000
2. FRC1 - частный критерий: усл. вероятность его признака сред...	Корреляция усл.отн частот с о...	0.105				0.704	1.000	0.826	100.000
2. FRC1 - частный критерий: усл. вероятность его признака сред...	Сумма усл.отн частот по при...					0.720	1.000	0.837	100.000
3. FRC2 - частный критерий: условная вероятность его признака...	Корреляция усл.отн частот с о...	0.104				0.704	1.000	0.826	100.000
3. FRC2 - частный критерий: условная вероятность его признака...	Сумма усл.отн частот по при...					0.720	1.000	0.837	100.000
4. INF1 - частный критерий: количество знаний по А.Харкевичу, в...	Семантический резонанс зна...	0.111				0.740	0.988	0.846	99.669
4. INF1 - частный критерий: количество знаний по А.Харкевичу, в...	Сумма знаний	0.004				0.792	1.000	0.884	100.000
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в...	Семантический резонанс зна...	0.111				0.740	0.988	0.846	99.669
5. INF2 - частный критерий: количество знаний по А.Харкевичу, в...	Сумма знаний	0.004				0.791	1.000	0.884	100.000
6. INF3 - частный критерий: Уинклеров. разности между факти...	Семантический резонанс зна...	0.174				0.734	1.000	0.847	100.000
6. INF3 - частный критерий: Уинклеров. разности между факти...	Сумма знаний	0.102				0.753	1.000	0.859	100.000
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.095				0.744	0.949	0.834	98.013
7. INF4 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.000				0.854	1.000	0.921	100.000
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Семантический резонанс зна...	0.095				0.744	0.948	0.834	98.013
8. INF5 - частный критерий: ROI (Return On Investment); вероятн...	Сумма знаний	0.000				0.854	1.000	0.921	100.000
9. INF6 - частный критерий: разн усл и безуслов вероятностей; вер...	Семантический резонанс зна...	0.102				0.735	1.000	0.847	100.000
9. INF6 - частный критерий: разн усл и безуслов вероятностей; вер...	Сумма знаний	0.003				0.777	1.000	0.874	100.000
10. INF7 - частный критерий: разн усл и безуслов вероятностей; ве...	Семантический резонанс зна...	0.102				0.735	1.000	0.847	100.000
10. INF7 - частный критерий: разн усл и безуслов вероятностей; ве...	Сумма знаний	0.003				0.777	1.000	0.874	100.000

Помощь по меркам достоверности | Помощь по частным распределениям | TP,TN,FP,FN | (TP-FP)/(TN-FN) | (TP-FP)/(TP+FP)100 | Задать интервал сглаживания

Рис. 2. Результаты верификации с информацией о достоверности СК-моделей по F-критерию Ван Ризбергена и L1- и L2-критериям проф. Е.В. Луценко

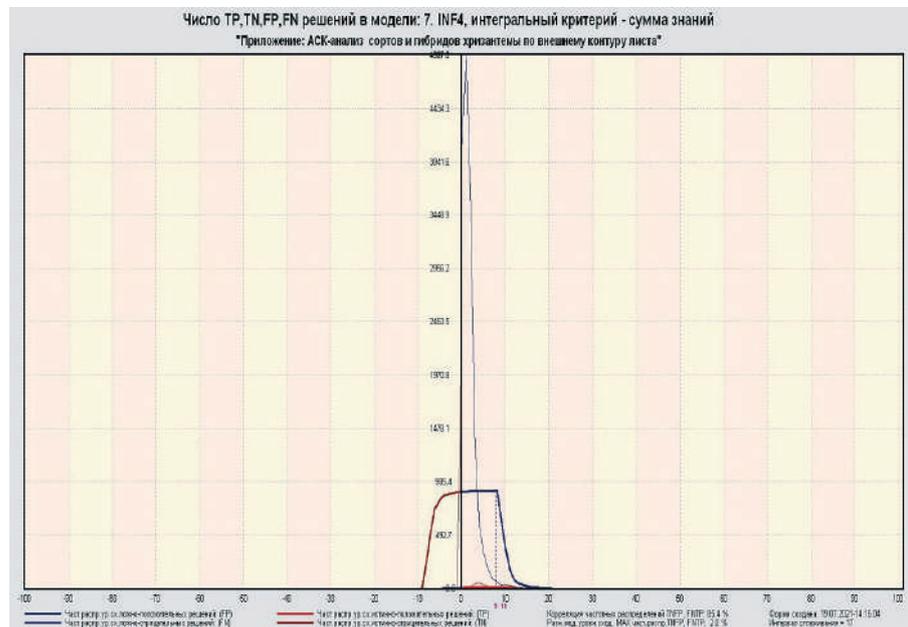


Рис. 3. Частотные распределения числа истинных и ложных положительных и отрицательных решений и их разности в СК-модели INF4 при интегральном критерии «Сумма знаний»

В интеллектуальной системе обучение модели происходит итерационно – пробуются различные модели, происходит переборка имеющихся параметров, сравниваются значения выбранной метрики и выбирается лучшая комбинация, поэтому необходимо принять метрику, которая объединяла бы в себе информацию о точности и полноте модели для упрощения принятия решения о том, какую модель запускать в качестве текущей. Именно такой метрикой в данном случае является классическая количественная F-мера достоверности моделей Ван Ризбергена [5], основанная на подсчёте суммарного количества верно и ошибочно классифицированных и неклассифицированных объектов обучающей выборки, и придающая одинаковый вес точности и полноте модели.

При малых выборках (до 400 объектов) в моделях преобладают низкие модули уровней сходства и их частотные распределения асимптотически приближаются к нормальным, поэтому при увеличении объёмов обучающей выборки достоверность модели уменьшается. Однако именно при этих выборках значения F-меры Ван Ризбергена близки к значениям мер L1 и L2. Так, модель INF3, как наиболее достоверная по показателю F-мера Ван Ризбергена ($F = 0,570$), даёт достоверность идентификации определения генотипа хризантемы по форме листа (контуру) 100,0 %, а достоверность неидентификации, т. е. определения генотипов к которым данный лист не относится – 62,1 %.

Для улучшения модели проводим вторую итерацию, исходя из полученного максимального значения достоверности модели INF3. Для этого в режиме 3.5 системы «Эйдос» (Синтез и верификация) модель INF1, заданная текущей по умолчанию, заменяется на INF3 и повторно изучается достоверность каждой частной модели в соответствии с этими мерами достоверности в режиме 3.4 (рис. 5).

Из рисунка 5 видно, что по критерию L1 проф. Е.В. Луценко [3, 5] наиболее достоверными являются СК-модели INF3 ($L1 = 0,775$) и INF4 ($L1 = 0,850$) с интегральным критерием «Сумма знаний» (при максимуме 1,000), что является очень хорошим результатом. Достоверность модели INF4 с интегральным критерием «Сумма знаний» по метрике L2 улучшилась с 0,921 до 0,936 (при максимуме 1,000), с более высоким показателем. Это подтверждает наличие в СК-модели достаточно сильной причинно-следственной зависимости между генотипическими особенностями хризантемы и количественными результатами описания внешнего контура листа.

В соответствии с информацией по достоверности моделей (рис. 5) модели INF4 присвоен статус текущей, и в ней проведена пакетная идентификация. Результаты распознавания отображаются в 12 формах, из которых мы приведём лишь две (рис. 6):

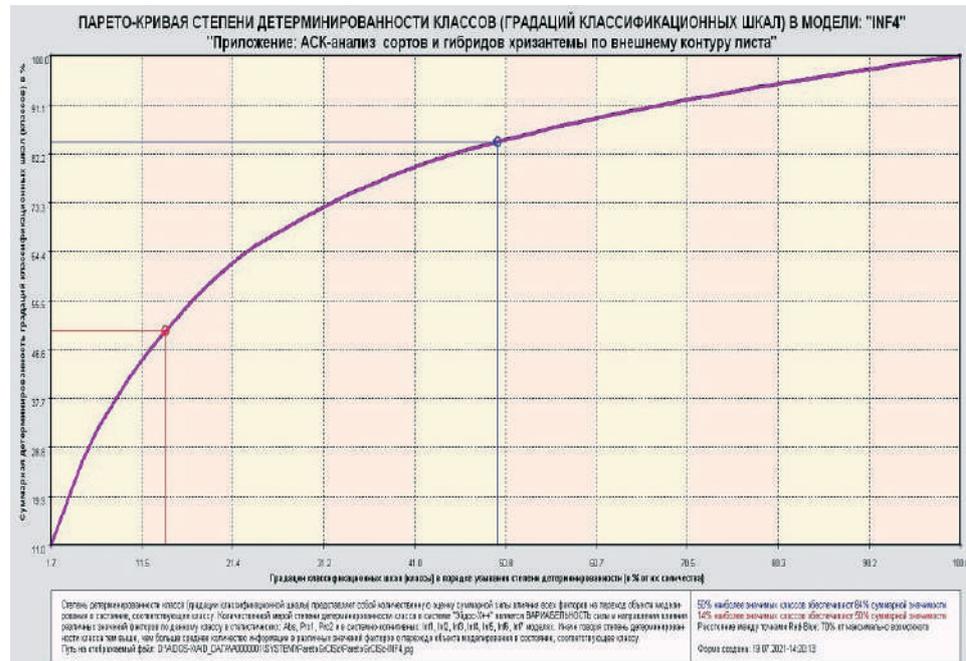


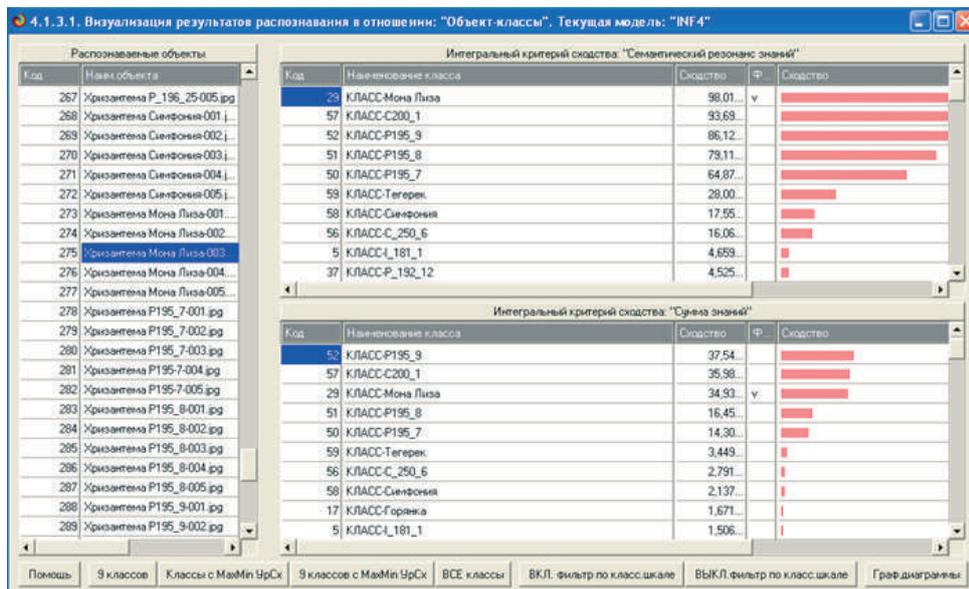
Рис. 4. Парето-кривая значимости влияния особенностей генотипа (класса)

3.4. Обобщ. форма по достов. моделей при разн. нит. крит. Текущая модель: "INF3"

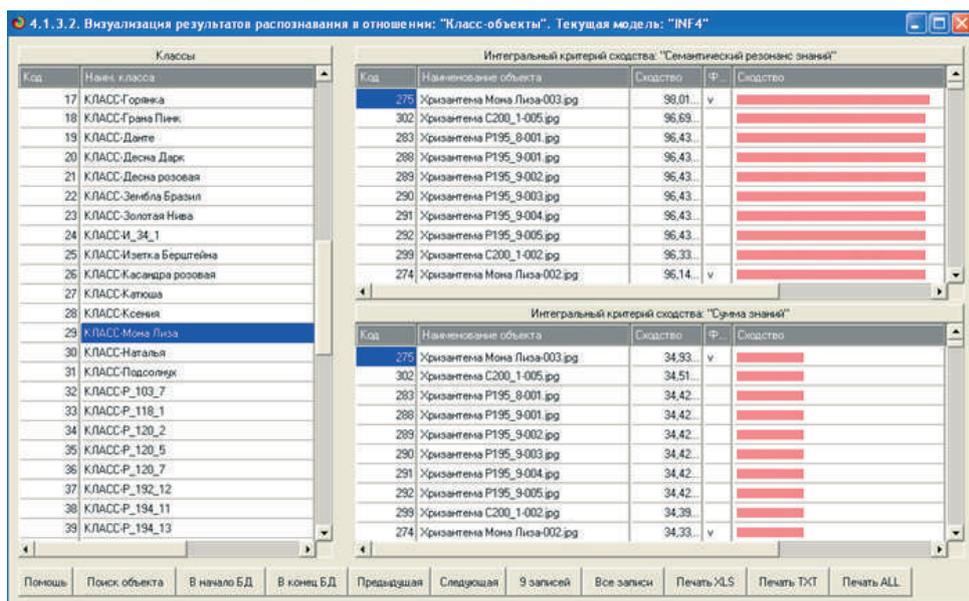
Наименование модели и частного критерия	Интегральный критерий	L1-мера пров. Е.В. Луценко	C	У	З	С	А-Точность модели AРископ = ATR/AATP+	A-Помота модели AРескал = ATR/AATP+	L2-мера пров. Е.В. Луценко	Процент процент
1. ABS - частный критерий: количество встреч соцветий: "клас...	Корреляция абс частот с обр...	0.648	-	-	-	-	0.655	1.000	0.792	100
1. ABS - частный критерий: количество встреч соцветий: "клас...	Сигна абс частот по признак...	0.597	-	-	-	-	0.675	1.000	0.806	100
2. PR1C1 - частный критерий: усл. вероятность i-го признака сред...	Корреляция усл отн частот с о...	0.648	-	-	-	-	0.656	1.000	0.792	100
2. PR1C1 - частный критерий: усл. вероятность i-го признака сред...	Сигна усл отн частот по приз...	0.597	-	-	-	-	0.675	1.000	0.806	100
3. PR1C2 - частный критерий: условная вероятность i-го признака...	Корреляция усл отн частот с о...	0.648	-	-	-	-	0.655	1.000	0.792	100
3. PR1C2 - частный критерий: условная вероятность i-го признака...	Сигна усл отн частот по приз...	0.597	-	-	-	-	0.675	1.000	0.806	100
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	0.760	-	-	-	-	0.711	0.988	0.827	99
4. INF1 - частный критерий: количество знаний по А.Харкевичу: в...	Сигна знаний	0.760	-	-	-	-	0.795	1.000	0.886	100
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Семантический резонанс зна...	0.760	-	-	-	-	0.711	0.988	0.827	99
5. INF2 - частный критерий: количество знаний по А.Харкевичу: в...	Сигна знаний	0.760	-	-	-	-	0.795	1.000	0.886	100
6. INF3 - частный критерий: Хинквэддат: разности между фактин...	Семантический резонанс зна...	0.747	-	-	-	-	0.691	1.000	0.817	100
6. INF3 - частный критерий: Хинквэддат: разности между фактин...	Сигна знаний	0.775	-	-	-	-	0.722	1.000	0.838	100
7. INF4 - частный критерий: ROI (Return On Investment): вероятно...	Семантический резонанс зна...	0.768	-	-	-	-	0.719	0.949	0.818	98
7. INF4 - частный критерий: ROI (Return On Investment): вероятно...	Сигна знаний	0.850	-	-	-	-	0.880	1.000	0.936	100
8. INF5 - частный критерий: ROI (Return On Investment): вероятно...	Семантический резонанс зна...	0.768	-	-	-	-	0.719	0.948	0.818	98
8. INF5 - частный критерий: ROI (Return On Investment): вероятно...	Сигна знаний	0.850	-	-	-	-	0.880	1.000	0.936	100
9. INF6 - частный критерий: разн усл и безуслов вероятностей, вер...	Семантический резонанс зна...	0.728	-	-	-	-	0.690	1.000	0.817	100
9. INF6 - частный критерий: разн усл и безуслов вероятностей, вер...	Сигна знаний	0.705	-	-	-	-	0.750	1.000	0.857	100
10. INF7 - частный критерий: разн усл и безуслов вероятностей, ве...	Семантический резонанс зна...	0.728	-	-	-	-	0.690	1.000	0.817	100
10. INF7 - частный критерий: разн усл и безуслов вероятностей, ве...	Сигна знаний	0.705	-	-	-	-	0.750	1.000	0.857	100

Помощь по мерам достоверности | Помощь по частным распределениям | TP; FN; FP; FN | (TP-FP); (TN-FN) | (T-F); (T+FP)100 | Задать интеграл сглаживания

Рис. 5. Экранная форма с информацией о достоверности моделей по F-критерию Ван Ризбергера и L1- и L2-критериям проф. Е.В. Луценко (2-я итерация)



A



B

Рис. 6. Экранная форма с результатами идентификации листьев с обобщенными образами листьев сортов и гибридов хризантемы:
А – объект-класс; Б – класс-объект

Символ «√» стоит напротив тех результатов прогнозирования, которые подтвердились на опыте, то есть соответствуют факту. Очевидно, что результаты идентификации довольно высокие (рис. 5).

Рассмотрим результаты кластерно-конструктивного анализа классов и признаков (когнитивные диаграммы и дендрограммы). В режиме 4.2.2.2 когнитивные диаграммы классов отражают сходство/различие классов.

На рисунке 7 приведена одна из выходных форм, количественно отражающая максимальную степень сходства образа конкретного сорта ('Mona Lisa White') с обобщёнными образами различных генотипов хризантемы. В кластер 'Mona Lisa White' вошли также гибриды, полученные от данного сорта, сорт голландского происхождения 'Tigerrag' и сорт российской селекции 'Симфония'. Все эти сорта и гибриды мелкоцветные, букетные. Форма и размер листьев коррелирует с размером соцветий. Знание данных корреляций может быть использовано при отборе на ранних стадиях селекции.

Информация о сходстве/различии классов, содержащаяся в матрице сходства, может быть визуализирована не только в форме когнитивных диаграмм, но и в форме агломеративных дендрограмм, полученных в результате когнитивной кластеризации [8] (рис. 8).

Из полученной дендрограммы мы видим, что все значения признаков образуют два чётко выраженных кластера, объединённых в полюса конструкта (показаны синим и красным цветами). Хорошо видна группировка сортов и гибридов хризантемы по детерминируемым ими количественным признакам контура листа. Значения признаков на полюсах конструкта признаков обуславливают переход объекта моделирования в состояния, соответствующие классам, представленным на полюсах конструкта классов (рис. 7).

Первый кластер (показан синим цветом на рисунке 8) включает преимущественно сорта голландского происхождения, имеющие крупные соцветия: 'Anastasia Green', 'Desna Dark', 'Baltica White', 'Bigoudi Red', 'Westland Regal', 'Wilhelmina', 'Zembla Brazil', 'Izetka Bernstein', 'Ksenia', 'Grand Pink', 'Natalia', 'Princess Armgard Bronze', P-196-27, P-201-1, P-196-26, 'Gagarin White', 'Ariana Lime', 'Dante', 'Горянка', 'Desna pink', P-30-24.

Гибриды P-192-12, P-194-13, полученные от 'Izetka Bernstein' в качестве материнского организма, оказались во втором кластере, что говорит о его слабой способности передавать хозяйственно-ценные признаки, несмотря на достаточно высокую продуктивность. Во второй кластер (красным цветом на рисунке 8) вошли мелкоцветные сорта и гибриды: 'Золотая Нива', 'Cassandra', 'Regina', 'Садко', 'Saratov', 'Tigerrag', 'Annecy White', 'Симфония', 'Etrusko', 'Mona Lisa White', P-120-5, С-200, И -34-5, P-192-12, P-194-13, P-196-15, P-196-25, P-195-7, P-195-8, P-195-9.

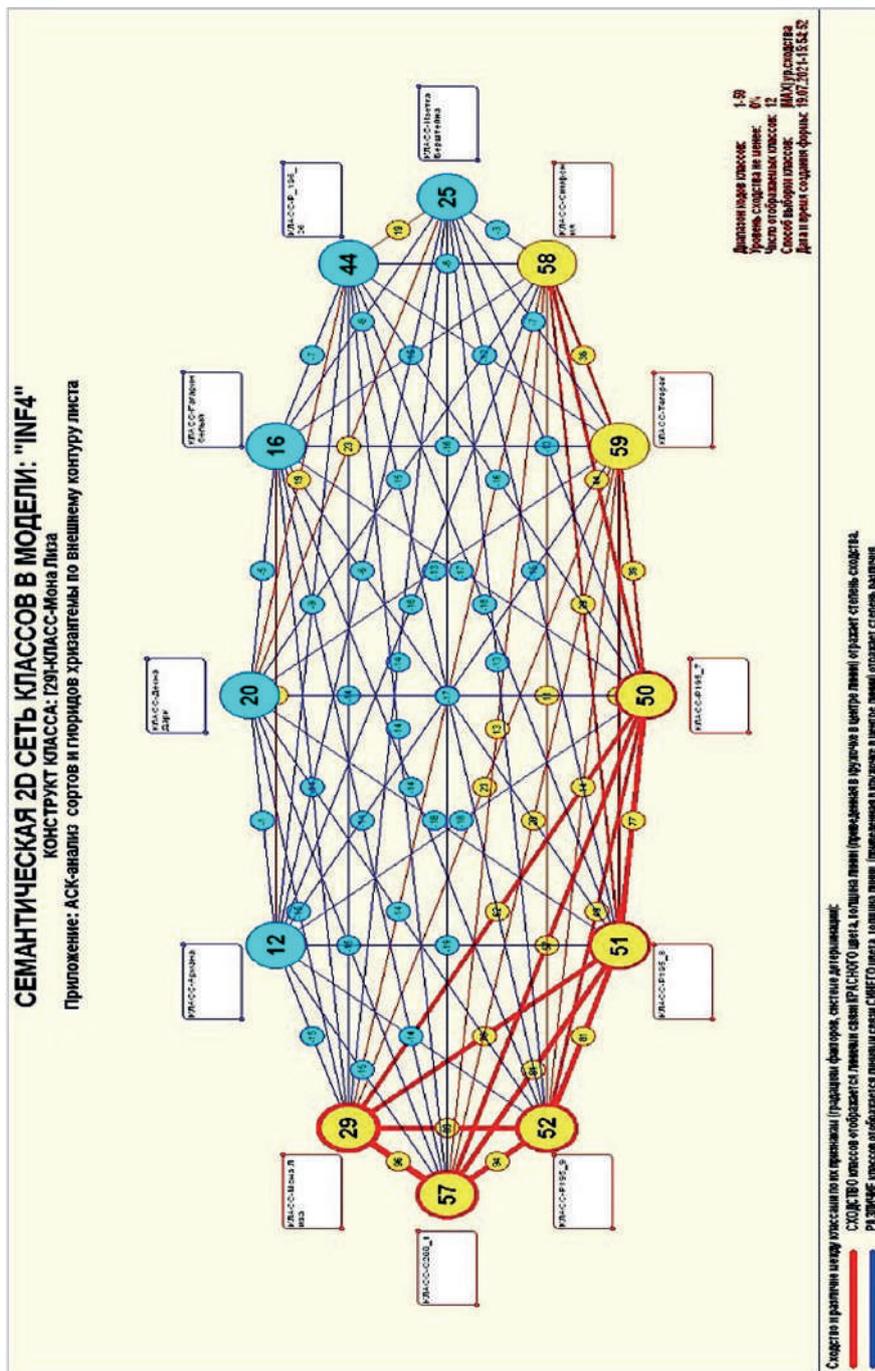


Рис. 7. Когнитивная диаграмма, отражающая степень сходства друг с другом обобщенных образов различных классов хризантемы по количественным признакам контура листа на примере конструкта сорта ('Mona Lisa White') с обобщенными образами различных генотипов хризантемы

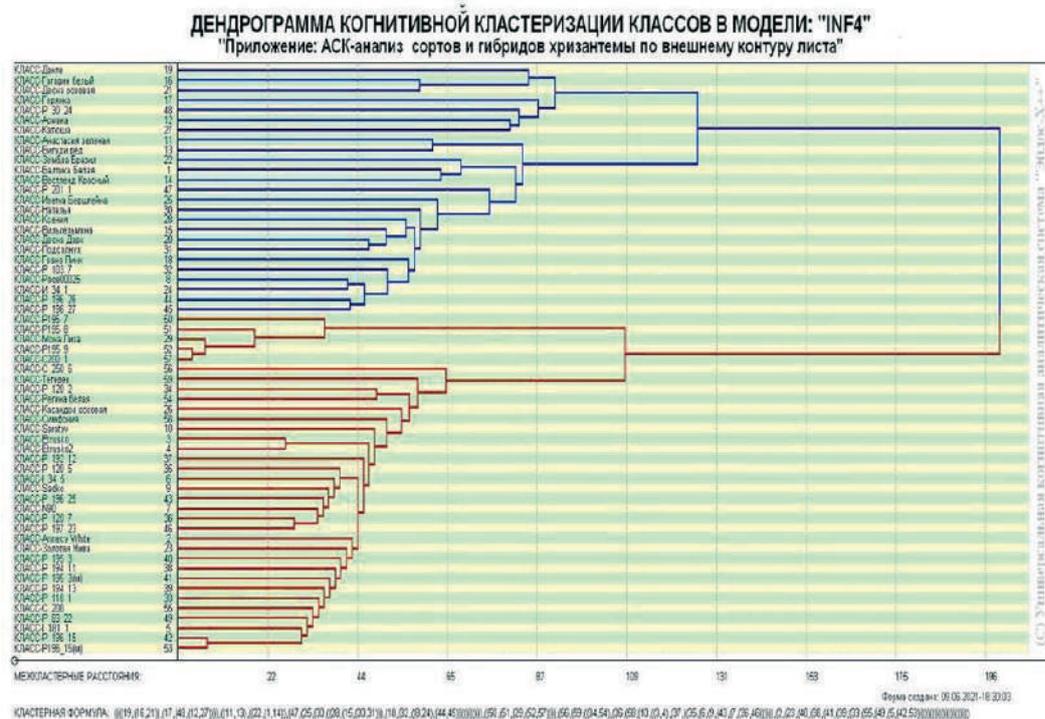


Рис. 8. Дендрограмма когнитивной агломеративной кластеризации классов, отражающая сходство/различие сортов и гибридов хризантемы по количественным признакам контура листа

Сравнение полученных результатов кластеризации АСК-анализом с фенотипической UPGMA-дендрограммой коллекции хризантем, основанной на 20 важных признаках сортов и гибридов, показывает сходные кластеры [18].

Хризантема садовая характеризуется высоким уровнем полиморфизма, что сказывается на фенотипической и генотипической изменчивости в гибридном потомстве. Например, гибриды P-196-15, P-196-25, P-195-7, P-195-8, P-195-9, C-200, полученные от ‘Mona Lisa White’ в качестве материнского растения, находятся в одном кластере. В то время как гибриды P-196-27, P-196-26 от этого же сорта оказались в разных кластерах. Это может служить основанием для предположения о большей возможности ‘Mona Lisa White’ передавать свои признаки потомству.

Выводы. Таким образом, по контурам листьев при помощи автоматизированного системно-когнитивного анализа (АСК-анализа), как интеллектуальной измерительной технологии, можно провести идентификацию различных сортов и гибридов представителей рода *Chrysanthemum*. Применение когнитивной агломеративной кластеризации позволило

разделить изучаемые генотипы хризантемы коллекции ФИЦ СЦ РАН на два основных кластера, включающие более мелкие группы. Данный анализ позволяет судить о генетических дистанциях между потомками растений и их родителями. АСК-анализ подтверждает сложную полиплоидную структуру хризантемы садовой, позволяет выявить значительный полиморфизм между родителями и потомками, а также между потомками одной комбинации скрещивания.

*Публикация подготовлена в рамках реализации
ГЗ ФИЦ СЦ РАН №№ 0492-2021-0008 и 0492-2021-0009*

Библиографический список

1. Гутиева Н.М. Методические подходы к отдалённой гибридизации рода *Pelargonium* // Современная методология, инструментарий оценки и отбора селекционного материала садовых культур и винограда: монография. – Краснодар: СКФНЦСВВ, 2017. – С. 261-268. – ISBN 978-5-98272-114-3.
2. Гутиева Н.М., Киселёва Н.С. Применение автоматизированного системно-когнитивного анализа в селекции пеларгонии // Субтропическое и декоративное садоводство. – 2018. – Вып. 67. – С. 73-82. – ISSN 2225-3068.
3. Луценко Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и её применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): монография. – Краснодар: КубГАУ, 2002. – 605 с.
4. Луценко Е.В. Метризация измерительных шкал различных типов и совместная сопоставимая количественная обработка разнородных факторов в системнокогнитивном анализе и системе «Эйдос» // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – 2013. – №08 (092). – eISSN: 1990-4665. – С. 859-883. – Режим доступа: <http://ej.kubagro.ru/2013/08/pdf/58.pdf>.
5. Луценко Е.В. Инвариантное относительно объёмов данных нечёткое мультиклассовое обобщение F-меры достоверности моделей Ван Ризбергера в АСК-анализе и системе «Эйдос» / Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2017. – № 02(126). – С. 1-32. – IDA [article ID]: 1261702001. – Режим доступа: <http://ej.kubagro.ru/2017/02/pdf/01.pdf>.
6. Луценко Е.В., Бандык Д.К. Автоматизированный системно-когнитивный анализ изображений по их внешним контурам (обобщение, абстрагирование, классификация и идентификация) // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – 2015. – № 110 (06). – С. 138-167. – eISSN: 1990-4665. – Режим доступа: <http://ej.kubagro.ru/2015/06/pdf/09.pdf>.
7. Луценко Е.В., Бандык Д.К., Трошин Л.П. Решение задач ампелографии с применением АСК-анализа изображений листьев по их внешним контурам (обобщение, абстрагирование, классификация и идентификация) // Политематический сетевой электронный научный журнал Куб ГАУ (Научный журнал КубГАУ) [Электронный ресурс]. – 2015. – № 112(08). – С. 862-910. – eISSN: 1990-4665. – Режим доступа: <http://ej.kubagro.ru/2015/08/pdf/64.pdf>

8. Луценко Е.В., Бандык Д.К., Трошин Л.П. Количественное измерение сходства-различия клонов винограда по контурам листьев с применением АСК-анализа и системы «Эйдос» // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – 2016. – № 116(02). – С. 1205-1228. – eISSN: 1990-4665. – Режим доступа: <http://ej.kubagro.ru/2016/02/pdf/77.pdf>
9. Луценко Е.В., Коржаков В.Е. Метод когнитивной кластеризации или кластеризация на основе знаний (кластеризация в системно-когнитивном анализе и интеллектуальной системе «Эйдос») // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – 2011. – № 07(071). – С. 528-576. – Режим доступа: <http://ej.kubagro.ru/2011/07/pdf/40.pdf>.
10. Методика государственного сортоиспытания сельскохозяйственных культур / под ред. С. Крылатова. – Вып. 6 (декоративные культуры). – М.: Колос, 1968. – 222 с.
11. Методика проведения испытаний на отличимость, однородность и стабильность хризантемы (многолетней) *Chrysanthemum* sp. // Официальный бюллетень Гос. комиссии РФ по испытанию и охране селекционных достижений. – 1995. – № 3. – 12 с.
12. Мохно В.С., Братухина Е.В., Якушина Л.Г. Селекция хризантемы в условиях влажных субтропиков Краснодарского края // Субтропическое и декоративное садоводство. – 2017. – Вып. 63. – С. 78-85. – ISSN 2225-3068.
13. Пидгайная Е.С., Репецкая А.И., Маркина Л.А., Решетникова Л.Ф. Методика сортооценки хризантемы садовой с учётом направления использования // Таврический вестник аграрной науки. – 2018. – № 3(15). – С. 88-97. – ISSN 2542-0720.
14. Программа Северо-Кавказского центра по селекции плодовых, цветочно-декоративных культур и винограда на период до 2030 г. / под общ. ред. Е.А. Егорова. – Краснодар: СКЗНИИСиВ, 2013. – 202 с. – ISBN 972-5-98272-096-2.
15. Рындин А.В., Мохно В.С. Методические подходы к созданию современных сортов садовых культур в субтропиках России // Субтропическое и декоративное садоводство. – 2012. – Вып. 47. – С. 111-117. – ISSN 2225-3068.
16. Рындин А.В., Слепченко Н.А. Генофонд цветочно-декоративных культур ФГБНУ ВНИИЦиСК: поддержание и пополнение коллекций // Научное обеспечение устойчивого развития плодового и декоративного садоводства: мат. междунар. науч.-практ. конф., посвященной 125-летию ВНИИЦиСК и 85-летию Ботанического сада «Дерево Дружбы», г. Сочи, 23-27 сентября 2019 г. – Сочи: ВНИИЦиСК, 2019. – С. 311-318. – ISBN: 978-5-904533-32-8.
17. Траутвейн К.С., Клемешова К.В. История изучения хризантемы (*Chrysanthemum* × *hortorum* Bailey) во Всероссийском научно-исследовательском институте цветоводства и субтропических культур // Субтропическое и декоративное садоводство. – 2018. – Вып. 64. – С. 26-33. – ISSN 2225-3068.
18. Samarina L.S., Malyarovskaya V.I., Reim S., Yakushina L.G., Koninskaya N.G., Klemeshova K.V., Shkhalakhova R.M., Matskiv A.O., Shurkina E.S., Gabueva T.Y., Slepchenko N.A., Ryndin A.V. Transferability of ISSR, SCoT and SSR-Markers for *Chrysanthemum* × *morifolium* Ramat and Genetic Relationships among Commercial Russian Cultivars // Plants. – 2021. – Vol. 10 – 1302. – <https://doi.org/10.3390/plants10071302>.

**THE POSSIBILITIES OF USING ASK ANALYSIS
AND THE INTELLIGENT EIDOS SYSTEM TO IDENTIFY
CULTIVARS AND HYBRIDS OF CHRYSANTHEMUM
BY THE LEAF CONTOUR**

Yakushina L.G., Kiseleva N.S., Slepchenko N.A.

*Federal Research Centre
the Subtropical Scientific Centre of the Russian Academy of Sciences,
Sochi, Russia, e-mail: vishnya584@yandex.ru*

An important task in chrysanthemum breeding is to identify various cultivars and hybrids, particularly by the leaf contours, since the contour shape of a particular leaf contains information about the cultivar: both about the genotype and the phenotype. Automated system-cognitive analysis (ASC-analysis) and its software tools, the intelligent system "Eidos" using scanned images of leaves, are proposed as tools for creating a measuring system and its application. The paper shows the results of a quantitative assessment of the external leaf contour in chrysanthemum cultivars and hybrids, on the basis of which generalized images of the studied genotype classes are formed according to the leaf shape; the influence of the values of their genotypic features is studied. Using the most reliable INF4 model, with the integral criterion "Sum of knowledge", the problems of quantitative identification of different cultivars and hybrids by leaf contours were solved. The studied chrysanthemum cultivar samples that belong to the collection of the FRC SSC of RAS are divided into two main clusters, including smaller groups, according to the characteristics of the leaf contour determined by them.

Key words: chrysanthemum, cultivar, hybrid, image system identification, leaf, genotype, similarity-difference, generalized images, "Eidos" system, ASK analysis.