

ПРОМЕЖУТОЧНЫЙ ОТЧЕТ О ВЫПОЛНЕНИИ НИР

«Исследование возможности использования нейронных сетей для автоматической классификации и поиска сходных изображений для промышленных образцов и товарных знаков»

Шифр темы: **5-ИТ-2020**

Сроки проведения НИР: 01.01.2020 – 31.12.2021



Научный руководитель НИР
заведующий отделением 13
Бирюков С.Е.

Ответственный исполнитель НИР
СНС отд. 31
Зонтов Ю.В.

Исполнители
Отдел 31
Отдел 119



Цель исследования

Установление возможности использования нейронных сетей и других методов машинного обучения для автоматической классификации и поиска сходных изображений промышленных образцов и товарных знаков

Задачи исследования

1 этап. Анализ требований ФИПС в отношении изображений промышленных образцов и товарных знаков, приводимых в документах по промышленной собственности (заявок, регистраций и прочих опубликованных материалов).

2 этап. Проведение экспериментов для исследования возможности использования нейронных сетей и других методов машинного обучения, а также анализа изображений, для автоматической классификации и поиска сходных изображений промышленных образцов и товарных знаков.

Актуальность исследования

Актуальность работы обусловлена необходимостью решения задачи автоматической классификации и поиска сходных изображений для промышленных образцов и товарных знаков, с учетом особенностей этих видов промышленной собственности.



- В настоящее время исследователями в области машинного обучения достигнуты значительные успехи в решении задачи автоматической классификации и поиска сходных изображений. Указанные технологии находят применение в различных сферах деятельности человека, в том числе в сфере охраны промышленной собственности.
- Задача классификации и сравнения изображений товарных знаков облегчается наличием Венской классификации изобразительных элементов знака, а так же тем, что, как правило, одному знаку соответствует одно характерное изображение.
- Аналогичная задача для промышленных образцов имеет ряд особенностей: образец может описываться множеством изображений, изображения могут иметь различные размеры, образец может описывать группу разнородных объектов или часть объекта и т.д.
- Актуальность работы обусловлена необходимостью решения задачи автоматической классификации и поиска сходных изображений для промышленных образцов и товарных знаков, с учетом особенностей этих видов промышленной собственности.






1 ЭТАП

01.01.2020 – 31.12.2020

- 
- 1) Составление реестра имеющихся изображений промышленных образцов и товарных знаков для тестовых целей;
 - 2) Анализ требований ФИПС в отношении изображений ПО и ТЗ и составление их классификации на базе полученного массива изображений;
 - 3) Анализ существующих подходов, методов и программных пакетов машинного обучения и анализа изображений;
 - 4) Выработка критериев схожести, пригодных для использования в методах машинного обучения;
 - 5) Выбор конкретных программных средств для проведения дальнейших экспериментов, как с точки зрения соответствия целям НИР, так и с точки зрения практической доступности

2 ЭТАП

01.01.2021 – 31.12.2021

- 
- 1) Освоение выбранных программных средств и подготовка обучающих и проверочных наборов данных;
 - 2) Проведение экспериментов по машинному обучению с выбранным программным обеспечением для исследования возможности его применения в целях данной работы
- 



Основные планируемые результаты работы:

1 этап - аналитическая справка о существующих подходах, методах, доступных программных пакетах машинного обучения и анализа изображений, а также критериях определения сходства изображений промышленных образцов и товарных знаков, приводимых в документах по промышленной собственности, с учетом требований ФИПС

2 этап - рекомендации по применению существующих подходов, методов и доступных программных пакетов машинного обучения и анализа изображений для автоматической классификации и поиска сходных изображений для промышленных образцов и товарных знаков, с учетом выработанных критериев

Результаты работы могут быть использованы:

При разработке новых информационных систем ФИПС

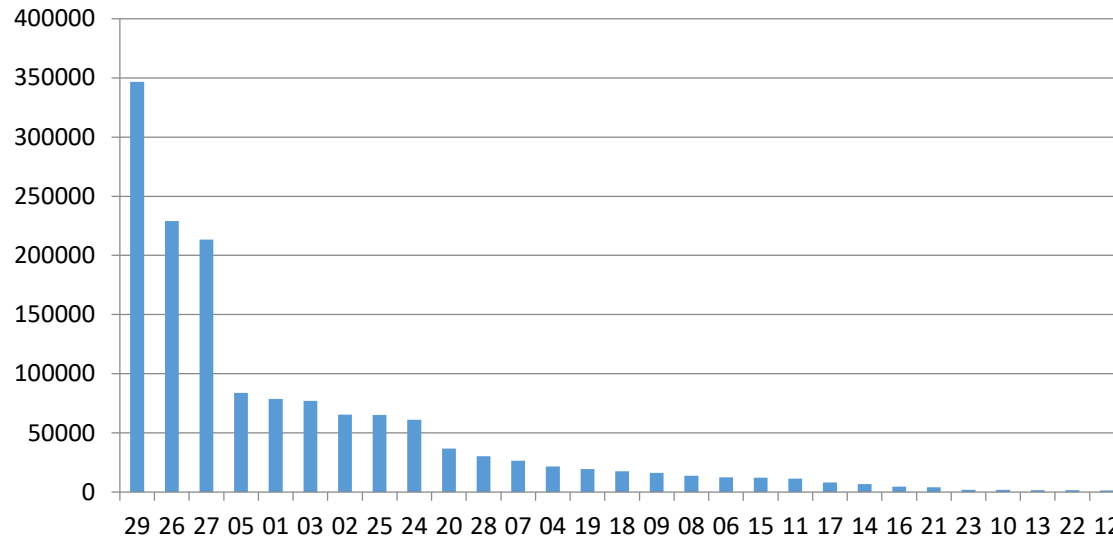
Потребители результатов исследования

Эксперты отделений экспертизы промышленных образцов и товарных знаков, разработчики информационных систем

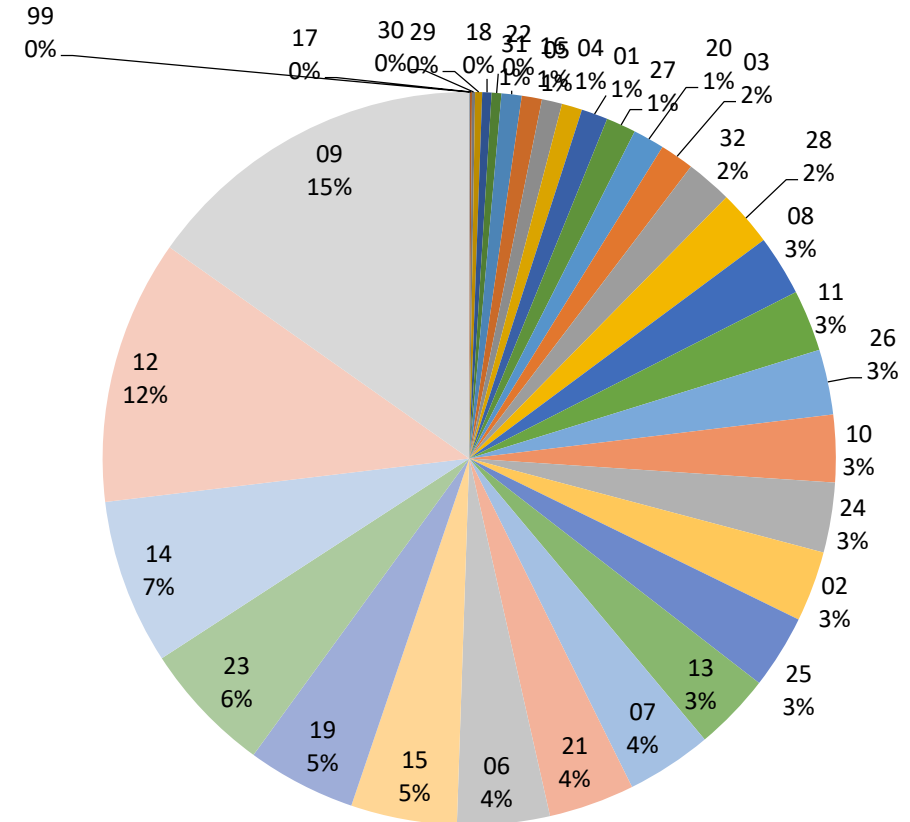


Составлен реестр имеющихся изображений промышленных образцов и товарных знаков для тестовых целей

Количество изображений по Венским классам



Количество изображений по классам МКПО



Рассмотрен состав изображений ПО и ТЗ в электронных хранилищах АСТЗ РФ и АСББД ЗПО

Новизна результата исследования заключается в использовании МКПО при анализе массивов изображений



★ Проведен анализ требований ФИПС в отношении изображений ПО и ТЗ и составление их классификации на базе полученного массива изображений

★ **Формализуемые признаки**

★ эстетические особенности внешнего вида изделия ПО (форма, конфигурация, орнамент, сочетание цветов, линий, контуры изделия, текстура или фактура материала изделия), а также изобразительные обозначения товарных знаков (изображения на плоскости живых существ, предметов, природных и иных объектов, композиции линий, пятен, любых фигур)

Трудно формализуемые процедуры

сравнение общих впечатлений, сравнительный анализ отличительных существенных признаков заявленного промышленного образца в целях проверки творческого характера существенных признаков

Новизна результата исследования заключается в изучении требований в контексте их применения в методах машинного обучения



★ Проведен анализ существующих подходов, методов и программных пакетов машинного обучения и анализа изображений

★ Для успешного решения задач поиска похожих изображений (для ТЗ и ПО) и выявления нарушения единства (для ПО) следует применять методы как *image classification* так и *object detection*, *image segmentation* и *scene recognition*, комбинируя последние с другими методами машинного обучения и эвристическими алгоритмами.

★ Возможно применение ансамблевых методов машинного обучения.

Для ПО важным является не только наличие тех или иных объектов на изображении, но и их взаимное расположение. В связи с этим представляется полезным применение методов *scene recognition*.

Новизна результата исследования заключается в изучении существующих методов машинного обучения в контексте их применения к изображениям ПО и ТЗ



- ★ *Выработаны критерии схожести, пригодные для использования в методах машинного обучения*
 - ★ *Для задачи поиска похожих изображений рекомендуется применение комбинированных подходов, при которых нейронные сети применяются для получения векторного представления изображения (“эмбеддинга”), а для их сравнения применяются другие методы машинного обучения.*
 - ★ *При сравнении “эмбеддингов” изображений требуется построение ранжированной выборки. Представляется целесообразным проводить сравнение внутри подмножества изображений того же класса, что и анализируемое изображение.*
- Для предсказания нарушения единства на основе существующих отказных решений строится модель классификации “похож - не похож”.*

Новизна результата исследования заключается в использовании комбинированного подхода.



Произведен выбор конкретных программных средств для проведения дальнейших экспериментов, как с точки зрения соответствия целям НИР, так и с точки зрения практической доступности



При выборе программных средств для дальнейших экспериментов предлагается применять следующие критерии



- Функциональные возможности
- Популярность
- Простота использования
- Открытый код
- Бесплатность

Были отобраны следующие продукты: **OpenCV, Tensorflow, PyTorch, LIRE**

Новизна результата исследования заключается в том, что критерии выбора программных средств направлены на соблюдение рекомендаций правительства РФ в части свободного ПО, а также минимизации затрат



Материалы, подготовленные в рамках 1 этапа НИР, составляют аналитическую справку о существующих подходах, методах, доступных программных пакетах машинного обучения и анализа изображений, а также критериях определения сходства изображений промышленных образцов и товарных знаков, приводимых в документах по промышленной собственности, с учетом требований ФИПС.

На втором этапе НИР в 2021 году планируется:

- Освоение выбранных программных средств и подготовка обучающих и проверочных наборов данных
- Проведение экспериментов по машинному обучению с выбранным программным обеспечением для исследования возможности его применения в целях данной работы
- Актуализация реестра изображений для учета материалов новых заявок
- Обогащение коллекции изображений за счет открытых данных EUIPO
- Подготовка рекомендаций по применению существующих подходов, методов и доступных программных пакетов машинного обучения и анализа изображений для автоматической классификации и поиска сходных изображений для промышленных образцов и товарных знаков, с учетом выработанных критериев



Спасибо за внимание!

