



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ БАЗЫ ДАННЫХ, ОХРАНЯЕМОЙ
АВТОРСКИМИ ПРАВАМИ

Номер регистрации (свидетельства):
2024620519

Дата регистрации: 01.02.2024

Номер и дата поступления заявки:
2023624869 15.12.2023

Дата публикации и номер бюллетеня:
01.02.2024 Бюл. № 2

Контактные реквизиты:
нет

Автор(ы):

Голубничий Артем Александрович (RU),
Яблонцева Арина Дмитриевна (RU)

Правообладатель(и):

Федеральное государственное бюджетное
образовательное учреждение высшего
образования "Хакасский государственный
университет им. Н.Ф. Катанова" (RU)

Название базы данных:
EMNIST TCyrillic

Реферат:

База данных содержит сведения о 74088 изображениях, относящихся к 147 символам, из них: 33 основные буквы кириллицы в двух регистрах, 29 букв расширенного кириллического алфавита в двух регистрах, 23 пунктуационных символа. Каждый символ представлен в 504 вариантах с разными шрифтами, а также различными степенями искажения: добавление шума, размытие (4 уровня), размытие со сдвигом (12 уровней), oilpaint, emboss и implode. База данных может быть использована для тренировки и обучения моделей машинного обучения, в том числе для построения эффективных алгоритмов распознавания символов тюркских языков. Тип ЭВМ: IBM PC-совмест. ПК. ОС: macOS 10.12 и выше, Windows 7 и выше, Linux-совмест. ОС (Ubuntu/Debian).

Вид и версия системы управления базой данных: Среда статистической обработки данных R версии не ниже 4.0

Объем базы данных: 135,1 МБ