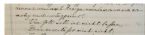# Decoding of handwritten diaries
A.V. Sukhovo-Kobylin

Morozov Ivan Dmitrievich, group 417
Supervisor: Doctor of Technical Sciences Leonid Moiseevich Mestetsky

Lomonosov Moscow State University
Faculty of Computational Mathematics and Cybernetics
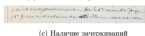Department of Mathematical Forecasting Methods

2024 г.

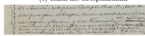# Recognition of manuscripts by A. V. Sukhovo-Kobylin
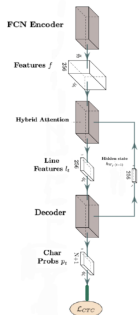

(a) Наличие иностранного языка

(b) Наличие надстрочных ветвей

(c) Наличие зачеркиваний

(d) Меняющаяся высота строк

FCN Encoder

Features f

Hybrid Attention

Line Features l_t

Decoder

Char Probe p_t

L_{CTC}

Архитектура VAN

The main goal:

- Transcription of manuscripts (≈ 10000 pages) with a Character Error Rate of less than 5%

An alternative goal:

- To develop an algorithm for highlighting keywords based on weak decoding

Key features of the solution:

1. **Data preprocessing.** Preparation and markup of the dataset.
2. **Using CTC-loss.** Use to work with gaps and unevenness.
3. **Vertical Attention Network architecture.** Analysis of handwritten pages with vertical attention.