

マルチモーダルモデルによる Few-Shot アイヌ語 OCR

安岡孝一 (京都大学)・安岡素子 (京都外国語大学)

はじめに

2026 年 2 月 16 日に発表^{a)}された Qwen3.5 は、マルチモーダルかつ多言語対応を謳っており、実際、日本語 OCR としても、そこそこの性能を発揮する。しかし、アイヌ語は読めない。多言語対応にアイヌ語が含まれておらず、アイヌ語の小書きカタカナ [1, 2] が読めないのだ。この問題に対し、われわれは、小書きカタカナのトークン追加をおこなった上で、Qwen3.5 に追加学習 (Fine-Tuning) をおこなう手法を提案した [3]。この手法は一定程度の成果は収めたものの、われわれの手元のマシンでは 9B (90 億) パラメータが追加学習の限界であり、それより大きなパラメータ数のモデルは試せなかった。

そうしたところ、Qwen3.5 が nvidia free inference endpoint で公開されているのを知った。397B パラメータ^{b)}が 7 月 27 日まで、122B パラメータ^{c)}が 7 月 20 日までの限定公開ではあったものの、user-assistant 型の API が使用可能である。ただし、この API では、

トークン追加も Fine-Tuning もできない。さて、どうするか。

Few-Shot アイヌ語 OCR

Fine-Tuning がダメなら、Few-Shot [4] はどうだろう。いくつかの例題を準備し (図 2)、各例題の画像を user prompt に、文字起こしテキストを assistant prompt に、順に配置する。その後に、OCR をおこないたい画像 (図 1) を user prompt として配置する。各 user prompt には、画像に加え「OCR Ainu sentences without translation.」という指示を付記しておく。

図 3 のプログラムで、Qwen3.5-397B-A17B に対し Three-Shot (図 2 の i・ii・iii を用いる) でアイヌ語 OCR を試したところ、以下の結果が得られた。

ウエランカラ

令和 7 年度アイヌ工芸品展「アイヌの美・彩りと輝き」アニ アンレヘ チコレ ワ、ウサムベカ イコロ チアマ ワ エチヌカレアン。

アイヌ アナクネ テエタ ワノ アンクル ネワ 20 イク ネ バ エムコ バクノ ヤウンモシリ、ヤンケモシリ (カラフト)、ルトム オッタ オカイ。タネ ヤウンモシッ タ カイ サモモシッ タ カイ、モシリ エビッタ オカイ。テエタ アナクネ チクニ、ニカフ、キナカフ、チコイキナ ボネ、キラウ、シキテ、ルシ、チエフカフ、セイ アニ ウサ オカイ ベカラ。ネアンベ オッタ ケシト アネイワンケ プ カイ アン。イノミ オッタ アネイワンケ プ カイ アン。ピシカン モシリ エオカイ ウタシ トウラ、アイヌ ウイマム ワ、センカキ、ヌイト、カネ、タマコロ ワ ネアンベ アニ ビリカ イカラカラ ビリカ イヌイェアイヌ キ。

タバン イキ オッタ イロホ ビリカ プ、ヌベキ ビリカ プ チヌムケ ワ チサンケ。イコロ エチヌカレアン カトツ エネアニ。センカキ トウラ アンニヌ ワ アンカラ イミカイ アン。センカキ トウラ アンチセ ワ アンカラ チタラベカイ アン。ルトム ウンクル イカラカラ ワ カラクツ カイ アン。ヤンケモシリ ウンクル コロ タマサイ カイアン ワ、ボンコンコ コロ タマサイ カイ アン。ウサ カネ ウサ ユッキラウ アニ アントウテ イコロ、イカヨビコロ、タンバクオフ、マキリ カイ アン。イノミ オッタ アネイワンケ プ トウキ、タカイサラ、パツチ カイ アン。

ネアンベ トウラノ クスロツ タシヌ チカッ 美恵子 (1948~2010) カラ ベ カイ エチヌカレアン。エエバキタ、貝澤徹 (イヌイェクル)、下倉洋之 (カネヌイェクル)、藤戸康平 (イコロカラル) タネ ウサ オカイ ビリカ プ カラ ワ コスル アシ ウタン ネ ワ クスネアンベ カイ チサンケ ワ エチヌカレアン。

タア アナクネ、アイヌプリ ビラサ クニ、公益財団法人アイヌ民族文化財団 (アヌコロトミ アニ アイヌプリ ビラサ ウタ) ケシ バキ ナネワ、タネ 30 スイ チキ シッタンナ。タンベ エチヌカラ ワ、テエタ アイヌプリ、タネ アン アイヌプリ、アムキリ クル インネ チキ ビリカ。イヨッタ イオシ、タバン イキ チコロ テッサマ、ビリカイコロ ウネルサ アウタラ ウサ ウサ ウンカスイ ウタラ、バセタラ チコヤイライケ シッタンナ。

北海道立釧路芸術館
京都府
京都府京都文化博物館
公益財団法人アイヌ民族文化財団

小書きの「プ」が出力されており、なかなか良い結果に見える。この結果と、図 1 に対する文字起こしテキスト (「プ」を 1 文字とみなし、連続する空白と改行を 1 文字にまとめた結果、総文字数 1124 字 [3]) とのレーベンシュタイン距離を見てみたところ、47 (単純誤り率 4.2%) だった。

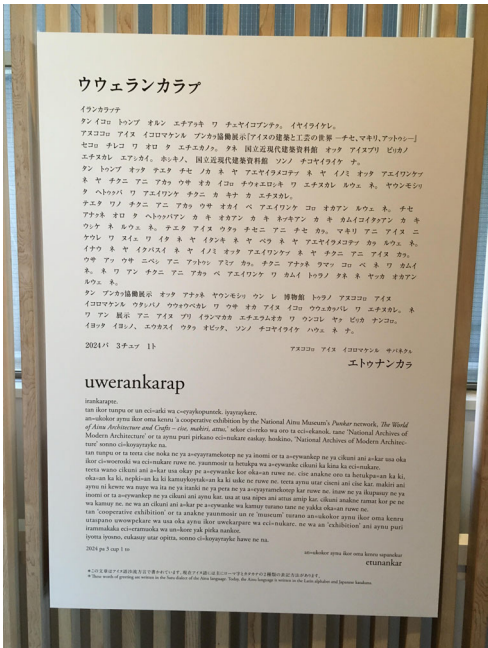
また、例題を Two-Shot (図 2 の i・ii) に減らした結果のレーベンシュタイン距離は 55 (単純誤り率 4.9%)、

図 1: 『アイヌの美・彩りと輝き』ウエランカラ

^{a)}<https://qwen.ai/blog?id=qwen3.5>

^{b)}<https://build.nvidia.com/qwen/qwen3.5-397b-a17b>

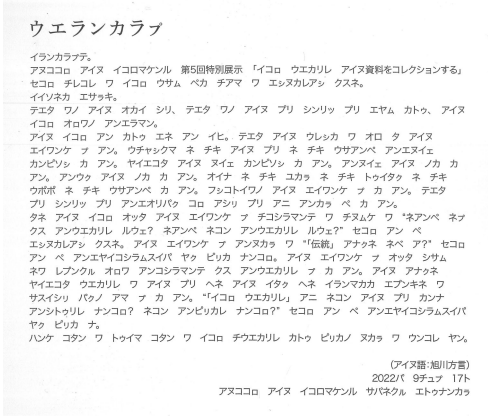
^{c)}<https://build.nvidia.com/qwen/qwen3.5-122b-a10b>



ウエランカラパ
イランカラパテ

タンイコロトウブンオルンエチアラキワチエヤイコブンテク。イヤイライケレ。
アヌコロアイヌイコロマケンルブンカラ協働展示『アイヌの建築と工芸の世界—チセ、マキリ、アットウシ—』セコロチレコ
ワオロタエチエカノク。タネ国立近代建築資料館オッタアイヌプリリカノエチヌカレエアンカ。ホシキ、国立近代建築
資料館ソノチコヤイライケナ。
タントウブンオッタテエチセノカネヤアエヤイラメコテプネヤイノミオッタアエイワングネヤチクニアニアカウウサ
オカイコロチウオエロシキワエチヌカレルウエネ。ヤウンモシリタヘトウクバワアエイワングネチクニカキナエチヌカレ。
テエタワノチクニアニアカウウサオカイベアエイワングネコロオカアンルウエネ。チセアナクネオロタヘトウクバアンカキ
オカアンカキネプキアンカカキムイコイタクアンカキウシケネルウエネ。テエタアイヌウタラチセニアニチセカラ。マキ
リアニアイヌニケウレワヌイエウイタネヤイタンキネヤペラネヤアエヤイラメコテプナラウエネ。イナウネヤイクバ
スイネヤイノミオッタアエイワングネヤチクニアニアイヌカラ。ウサアッウサニベシアニアットウシアミアカラ。チク
ニアナクネラマッコロベネワカムイネ。ネワアンチクニアニアカウベアエイワングネワカムイトウラノタネネヤッウオカア
ンルウエネ。
タンブンカラ協働展示オッタアナクネヤウンモシリウンレ博物館トウラノアヌコロアイヌイコロマケンルウタシノウウ
ウベカレワウサオカイアイヌイコロウエカラバレワエチヌカレ。ネワアン展示アニアイヌプリイランマカエチエラムオカ
ワウンコレイキヤリカナンコロ。
イヨッタイヨシハ、エウカアンイウタロオビッタ、ソノチコヤイライケハウエネナ。
2024バ3チュブ1ト
アヌコロアイヌイコロマケンル
サバネクル
エトウナンカラ
uwerankarap.
irankarapte.
tan ikor tuntu or un eci=arki wa c=eayakopuntek. iyayarkere.
an=ukokor aynu ikor oma kenru 'a cooperative exhibition by the National Ainu Museum's Punkar network, The World of Ainu Architecture
and Crafts — cise, makiri, attos, ' sekor ci=reko wa oro ta eci=ekanok. tane ' National Archives of Modern Architecture ' or ta aynu puri
pirkano eci=nukare easkay, huskino, ' National Archives of Modern Architecture ' sono ci=koyayrayke na.
tan tuntu or ta teeta cise noka ne ya a=eayaraketep ne ya inomi or ta a=eyanwape ne ya cikuni ani a=kar usa oka ikor ci=woeroski wa
eci=nukare ruwe ne. yaunmosir ta hetukpa wa a=eywanke cikuni ka kina ka eci=nukare.
teeta wano cikuni ani a=kar usa okay pe a=eywanke kor oka=an ruwe ne. cise anakne oro ta hetukpa=an ka ki, oka=an ka ki, nepki=an ka ki
kamuykoytak=an ka ki uske ne ruwe ne. teeta aynu utar ciseni ani cise kar. makiri ani aynu ni kewe wa nuye wa ita ne ya itanki ne ya pera ne
ya a=eayaraketep kar ruwe ne. inaw ne ya ikupasuy ne ya inomi or ta a=eyanwape ne ya cikuni ani aynu kar. usa at usa nipes ani attus amip
kar. cikuni anakne ramat kor pe ne wa kamuy ne. ne wa an cikuni ani a=kar pe a=eywanke wa kamuy turano tane ne yakka oka=an ruwe ne.
tan 'cooperative exhibition' or ta anakne yaunmosir un re 'museum' turano an=ukokor aynu ikor oma kenru utaspano uwowepkare wa usa
oka aynu ikor uwepkare wa eci=nukare. ne wa an 'exhibition' ani aynu puri irammakaka eci=eramuka wa un=kore yak pirka nankor.
iyotta iynosno, eukasyu utar opitta, sono ci=koyayrayke hawe ne na.
2024 pa 3 cup 1 to
an=ukokor aynu ikor oma kenru sapanekur
etunankar

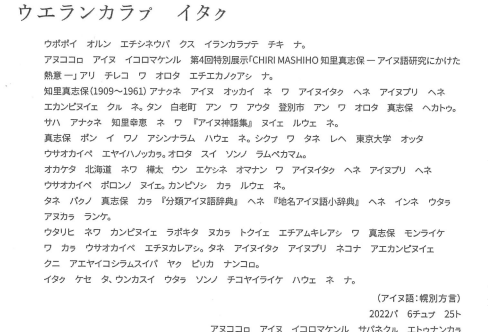
i) 1st-Shot 画像 (2024-03-01.jpg) と文字起こしテキスト



ウエランカラパ
ウエランカラパテ

イランカラパテ。
アヌコロアイヌイコロマケンル 第5回特別展示「イコロウエカリエアイヌ資料をコレクションする」セコロチレコレワイコロ
ウサムベカチアマワエシヌカレアシクスネ。
イイソネカエサキ。
テエタワノアイヌオカイシリ、テエタワノアイヌプリシリシリプリエヤムカトウ、アイヌイコロオロワノアンエラムン。
アイヌイコロアンカトウエネアンイヒ。テエタアイヌウレシカワオロタアイヌエイワングネアン。ウチヤシクマネチキアイ
ヌプリネチキウサアンベアヌエヌアイエカンビソシカアン。ヤイエコタアイヌヌイエカンビソシカアン。アンヌアイヌノカ
カアン。アンウクアイヌノカカアン。オйнаネチキユカラネチキトウイタクネチキウボボネチキウサアンベカアン。フシコ
トイワノアイヌエイワングネカアン。テエタシリシリシリプリアンエオリバコロアシリプリアニアンカラベカアン。
タネアイヌイコロオッタアイヌエイワングネアチコシラムンチワチヌムケワ「ネアンベネクスアンウエカリエルウエ?ネアン
ベネコンアンウエカリエルウエ?」セコロアンベエシヌカレアシクスネ。アイヌエイワングネアツツカラワ「伝統」アナクネ
ネベア?」セコロアンベアンエヤイコシラムスイバヤクピリカナンコロ。アイヌエイワングネアツツカラワ「伝統」アナクネ
ネワアンシラムンチクスアンウエカリエルウエ?」セコロアンベエヤイコシラムスイバヤクピリカナンコロ。アイヌエイワングネ
ハネイランマカカエプキネワサスイシリバノアマツカアン。「イコロウエカリエ」アニネコンアイヌプリカンナアンシ
トウリレナノコロ?ネコンアンビリカレナノコロ?」セコロアンベエヤイコシラムスイバヤクピリカナンコロ。
ハンケコタンワトウイマコタンワイコロチウエカリエカトウピリカノヌカラワウンコレヤン。
(アイヌ語: 旭川方言)
2022バ9チュブ1ト
アヌコロアイヌイコロマケンルサバネクルエトウナンカラ

iii) 3rd-Shot 画像 (2022-06-25.jpg) と文字起こしテキスト



ウエランカラパ
ウエランカラパテ

ウボボイオルンエチシネウバクスイランカラパテチキナ。
アヌコロアイヌイコロマケンル 第4回特別展示「CHIRI MASHIHO 知里真志保—アイヌ語研究にかけた熱意—」アリチレコ
ワオロタエチエカノクアシナ。
知里真志保(1909~1961)アナクネアイヌオッカネワアイヌイタクヘネアイヌプリヘネエカンビヌイェクルネ。タン白老町
アンワアウタ登別市アンワオロタ真志保ヘカトウ。サハアナクネ知里幸恵ネワ「アイヌ神謡集」ヌイェルウエネ。
真志保ボインワノアンシラムハウエネ。シクアワタネレハ東京大学オッタウサオカイベエヤイハノツカラ。オロタスイソ
ノラムベカマ。
オカケタ北海道ネワ樺太ウンエケシネオマナンワアイヌイタクヘネアイヌプリヘネウサオカイベボロンヌイェ。カンビソ
シカラルウエネ。
タネバクノ真志保カラ「分類アイヌ語辞典」ヘネ「地名アイヌ語小辞典」ヘネインネウタラアヌカラランケ。
ウタヒネワカンビヌイェラボキタヌカラトクイェエチアムキレアシワ真志保モナライケワカラウサオカイベエチヌカレアシ。
タネアイヌイタクアイヌプリネコナアエカンビヌイェクニアエヤイコシラムスイバヤクピリカナンコロ。
イタケセタ、ウンカアンイウタラソノチコヤイライケハウエネナ。
(アイヌ語: 樺列方言)
2022バ6チュブ25ト
アヌコロアイヌイコロマケンルサバネクルエトウナンカラ

図 2: Few-Shot 学習用アイヌ語画像と文字起こしテキスト

図 3: Three-Shot アイヌ語 OCR プログラム

でも有効かどうか、評価をおこなった。表3に、評価結果を示す。多くのモデルで有効性は示しているものの、gemma-4-31B-itはうまくいっていない。また、PaddleOCRやGLM-OCRでは、Few-Shotで出力が無限ループになるなど、本稿の手法が適用できなかった。

おわりに

Qwen3.5の各モデルに対して、Few-Shotでアイヌ語OCRをおこなう手法を提案し、その効果を示した。また、この手法を他のマルチモーダルモデルにも適用したところ、多くのモデルで効果が認められた。図3のプログラムは、様々なinference endpointで使える設計なので、ぜひ試してみてください。

参考文献

- [1] 佐藤知己: アイヌ語を記述するのに必要な文字セットについて, JIS 符号化文字集合調査研究委員会第 2 分科会 (WG2) 資料, JCS-2-8-02 (1996 年 11 月 25 日).
- [2] Addition of forty eight characters, JTC1/SC2/WG2 N2092 (September 13, 1999).
- [3] 安岡孝一, 安岡素子: Qwen3.5 によるカタカナアイヌ語 OCR の開発, 情報処理学会研究報告, Vol.2026-CH-141 『人文科学とコンピュータ』, No.3 (2026 年 5 月 16 日), pp.1-5.
- [4] Mohamed Ali Souibgui, Alicia Fornés, Yousri Kessentini, Beáta Megyesi: Few shots are all you need: A progressive learning approach for low resource handwritten text recognition, Pattern Recognition Letters, Vol.160 (August 2022), pp.43-49.

2026 Information Processing Society of Japan

表 1: Qwen3.5 各モデルにおいて図 1 を Few-Shot で OCR した際の評価

	Zero-Shot	One-Shot	Two-Shot	Three-Shot
Qwen/Qwen3.5-397B-A17B	114 (10.1%)	52 (4.6%)	55 (4.9%)	47 (4.2%)
Qwen/Qwen3.5-122B-A10B	115 (10.2%)	81 (7.2%)	59 (5.2%)	42 (3.7%)
Qwen/Qwen3.5-35B-A3B	228 (20.3%)	106 (9.4%)	97 (8.6%)	119 (10.6%)
Qwen/Qwen3.5-27B	108 (9.6%)	134 (11.9%)	63 (5.6%)	68 (6.0%)
Qwen/Qwen3.5-9B	217 (19.3%)	106 (9.4%)	80 (7.1%)	86 (7.7%)
Qwen/Qwen3.5-4B	126 (11.2%)	103 (9.2%)	91 (8.1%)	84 (7.4%)
Qwen/Qwen3.5-2B	122 (10.9%)	157 (14.0%)	90 (8.0%)	110 (9.8%)
Qwen/Qwen3.5-0.8B	412 (36.7%)	401 (35.7%)	361 (32.1%)	178 (15.8%)

※レーベンシュタイン距離 (単純誤り率)

表 2: AinuOCR 各モデル [3] において図 1 を Few-Shot で OCR した際の評価

	Zero-Shot	One-Shot	Two-Shot	Three-Shot
KoichiYasuoka/Qwen3.5-9B-AinuOCR	89 (7.9%)	102 (9.1%)	112 (10.0%)	113 (10.1%)
KoichiYasuoka/Qwen3.5-4B-AinuOCR	88 (7.8%)	83 (7.4%)	67 (6.0%)	72 (6.4%)
KoichiYasuoka/Qwen3.5-2B-AinuOCR	88 (7.8%)	73 (6.5%)	78 (6.9%)	77 (6.9%)
KoichiYasuoka/Qwen3.5-0.8B-AinuOCR	144 (12.8%)	452 (40.2%)	182 (16.2%)	167 (14.9%)

※レーベンシュタイン距離 (単純誤り率)

表 3: 他のマルチモーダルモデルにおいて図 1 を Few-Shot で OCR した際の評価

	Zero-Shot	One-Shot	Two-Shot	Three-Shot
Qwen/Qwen3.6-35B-A3B	107 (9.5%)	78 (6.9%)	62 (5.5%)	64 (5.7%)
Qwen/Qwen3.6-27B	260 (23.1%)	143 (12.7%)	103 (9.2%)	110 (9.8%)
moonshotai/Kimi-K2.6	78 (6.9%)	52 (4.6%)	31 (2.8%)	74 (6.6%)
moonshotai/Kimi-K2.7-Code	86 (7.7%)	127 (11.3%)	49 (4.4%)	46 (4.1%)
mistralai/Mistral-Medium-3.5-128B	119 (10.6%)	192 (17.1%)	71 (6.3%)	72 (6.4%)
mistralai/Mistral-Small-4-119B-2603	228 (20.3%)	208 (18.5%)	154 (13.7%)	139 (12.4%)
nvidia/nemotron-3-nano-omni-30b-a3b-reasoning	281 (25.0%)	70 (6.2%)	42 (3.7%)	52 (4.6%)
google/gemma-4-31B-it	181 (16.1%)	327 (29.1%)	208 (18.5%)	194 (17.3%)
PaddlePaddle/PaddleOCR-VL-1.5	288 (25.6%)	×	×	×
KoichiYasuoka/PaddleOCR-VL-1.5-Ainu	51 (4.5%)	×	×	×
PaddlePaddle/PaddleOCR-VL-1.6	112 (10.0%)	×	×	×
KoichiYasuoka/PaddleOCR-VL-1.6-Ainu	62 (5.5%)	×	×	×
zai-org/GLM-OCR	216 (19.2%)	×	×	×

※「×」は単純誤り率が 50%超