



Klasifikasi Tulisan Tangan Huruf Hijaiyah Anak Usia 6-8 Tahun Menggunakan Metode Support Vector Machine

Ahmad Maulidi Roofiad^{1*}, Cecep Nurul Alam¹, Aldy Rialdy Atmadja¹

¹Fakultas Sains dan Teknologi, UIN Sunan Gunung Djati Bandung

*Corresponding Author's e-mail: didinaja1787@gmail.com

Article History:

Received: November 25, 2025

Revised: December 12, 2025

Accepted: December 17, 2025

Keywords:

Hijaiyah

Image Classification Early

Childhood Support Vector

Machine

Abstract: This study aims to develop a handwritten Hijaiyah letter classification system for children aged 6–8 years using the Support Vector Machine (SVM) algorithm. The main problem in elementary education is the difficulty children face in recognizing and writing Hijaiyah letters due to the similarity of their shapes and variations in handwriting. The research process uses the CRISP-DM stages, consisting of problem understanding, data collection and preparation, modeling with SVM (GridSearch for hyperparameter tuning), and evaluation using a confusion matrix and f1-score. The dataset used consists of 2,100 images of handwritten letters from elementary school students. The results show that the SVM model with RBF kernel, $C=10$, and $\gamma=\text{"scale"}$ achieved the highest accuracy of 83.57%. This study demonstrates that an SVM-based machine learning approach can assist in recognizing Hijaiyah letters, making it a practical solution for teachers in teaching Hijaiyah writing.

Copyright © 2025, The Author(s).

This is an open access article under the CC-BY-SA license



How to cite: Roofiad, A. M., Alam, C. N., & Atmadja, A. R. (2025). Klasifikasi Tulisan Tangan Huruf Hijaiyah Anak Usia 6-8 Tahun Menggunakan Metode Support Vector Machine. *SENTRI: Jurnal Riset Ilmiah*, 4(12), 3762–3769. <https://doi.org/10.55681/sentri.v4i12.5077>

PENDAHULUAN

Pendidikan anak usia dini merupakan fondasi penting dalam penguasaan keterampilan dasar, termasuk membaca dan menulis huruf hijaiyah. Anak usia 6–8 tahun sering kesulitan mengenali dan menulis huruf hijaiyah akibat variasi tulisan tangan yang belum stabil. Hal ini sering menyebabkan kebingungan pada bentuk huruf yang mirip seperti 'ب', 'ت', dan 'ث' (Wahyuni, 2020; Yuliani & Surya, 2019). Di SDN 169 Pelita, jumlah siswa yang banyak menyulitkan guru dalam memberikan perhatian individual, sehingga berdampak pada penguasaan huruf hijaiyah yang kurang optimal.

Teknologi pengenalan pola menawarkan solusi efektif untuk masalah ini. Algoritma Support Vector Machine (SVM) dikenal handal dalam pengenalan pola data kompleks, termasuk tulisan tangan bervariasi. Penerapan SVM dalam *pattern recognition* telah terbukti efektif di berbagai aplikasi. Sebagai contoh, penelitian oleh Darwis et al. (2020) menunjukkan efektivitas Support Vector Machine dalam analisis sentimen pada data Twitter Komisi Pemberantasan Korupsi Republik Indonesia. Dalam studi tersebut, SVM, dikombinasikan dengan ekstraksi fitur TF-IDF, berhasil mengklasifikasikan opini masyarakat dengan hasil kelas positif 8%, negatif 77%, dan netral 15%, serta mencapai akurasi sebesar 82%, presisi 90%, *recall* 88%, dan f1-score

89% [11].

Walaupun demikian, eksplorasi penerapan SVM dalam pengenalan tulisan tangan huruf hijaiyah anak-anak masih terbatas, mengingat keunikan visual huruf hijaiyah dan karakteristik tulisan tangan anak usia 6–8 tahun yang kurang konsisten (Rosyid & Arifin, 2021; Arifin et al., 2023). Penelitian terkait pengenalan huruf hijaiyah menggunakan metode lain telah dilakukan. Galih Abdul Aziz (2023) melakukan penelitian di UIN Sunan Gunung Djati Bandung dengan menggunakan metode Convolutional Neural Network (CNN) pada 672 citra data latih dan 168 citra data uji tulisan tangan huruf hijaiyah anak usia 4–5 tahun di Maqdis, dan berhasil mencapai akurasi sebesar 88%. Aziz & Lestari (2023) menggabungkan CNN dan XGBoost untuk klasifikasi pengucapan huruf hijaiyah berdasarkan sanad dengan menggunakan 21.600 rekaman audio dari 40 responden, dibagi menjadi 16.200 data latih dan 5.400 data uji, dan melaporkan akurasi rata-rata 73,79% per huruf serta peningkatan F1-score setelah penambahan data. Fawwaz Ijlal Muqsith dkk. (2025) menerapkan CNN dengan fitur mel-spectrogram pada 8.904 sampel suara dan mencapai akurasi rata-rata 83,8% melalui aplikasi Android. Untuk domain citra, DE Putri dkk. (2024) menggunakan transfer learning EfficientNetB1 pada 2.800 citra bahasa isyarat huruf hijaiyah dan memperoleh akurasi lebih dari 95%. Selain itu, Siti Ruhu Salamah dkk. (2025) menerapkan CNN untuk mengklasifikasi sifat huruf hijaiyah, menunjukkan potensi metode deep learning tidak hanya pada bentuk visual huruf tetapi juga karakteristik fonetiknya. Studi-studi ini menegaskan relevansi penerapan metode machine learning, khususnya CNN dan model hibrida, dalam domain pengenalan huruf hijaiyah baik pada tulisan tangan, suara, maupun isyarat, sehingga dapat menjadi pembanding yang penting dengan efektivitas Support Vector Machine (SVM) dalam konteks tulisan tangan anak. Oleh karena itu, eksplorasi lebih lanjut terhadap SVM pada pengenalan tulisan tangan huruf hijaiyah anak usia 6–8 tahun menjadi signifikan untuk mengisi celah penelitian yang ada.

Penelitian ini bertujuan mengimplementasikan model berbasis Python menggunakan algoritma SVM untuk pengenalan pola tulisan tangan huruf hijaiyah anak usia 6–8 tahun di SDN 169 Pelita. Model ini kemudian diterapkan pada aplikasi sederhana berbasis mobile yang dikembangkan menggunakan React Native. Diharapkan model dan aplikasi ini dapat menjadi alat bantu pembelajaran yang adaptif, memberikan umpan balik langsung, dan meningkatkan pemahaman anak terhadap huruf hijaiyah, sekaligus berkontribusi pada pengembangan teknologi pendidikan berbasis machine learning.

METODE PENELITIAN



Gambar 1. *CRISP-DM*

Setiap tahap dilaksanakan secara sistematis agar dapat menghasilkan model klasifikasi tulisan tangan huruf hijaiyah pada anak usia 6–8 tahun yang akurat dan mudah diimplementasikan dalam pembelajaran.

1. Business Understanding, dilakukan dengan menganalisis kebutuhan pembelajaran huruf hijaiyah di lingkungan sekolah dasar, khususnya pada siswa usia dini yang sering mengalami kesulitan dalam membedakan bentuk huruf. Analisis kebutuhan didasarkan pada observasi proses belajar serta diskusi dengan guru untuk mengetahui kendala yang dihadapi dalam pembelajaran konvensional.
2. Data Understanding dilakukan dengan mengumpulkan dan mengeksplorasi data berupa 1.680 gambar tulisan tangan huruf hijaiyah yang dikumpulkan dari siswa kelas 1 dan 2 SD. Analisis awal dilakukan untuk memahami karakteristik, sebaran, serta keberagaman data agar model yang dihasilkan dapat merepresentasikan variasi tulisan tangan secara optimal.
3. Data Preparation, data gambar diproses melalui beberapa langkah, yaitu resize ke ukuran 64x64 piksel, konversi menjadi citra grayscale, penerapan Gaussian blur untuk mengurangi noise, serta adaptive thresholding dan cropping untuk menonjolkan area huruf.



Gambar 2. Tahapan *Preparation Data*

Setelah preprocessing, setiap gambar diekstraksi fiturnya menggunakan metode Histogram of Oriented Gradients (HOG), yang mampu menangkap karakteristik visual utama dari setiap huruf hijaiyah.

4. Modeling dilaksanakan dengan membangun model klasifikasi menggunakan algoritma Support Vector Machine (SVM). Pemilihan parameter model dilakukan dengan tuning hyperparameter (C, gamma, kernel) melalui teknik GridSearchCV dan validasi silang 5-fold, guna memperoleh kombinasi parameter terbaik yang memberikan akurasi tertinggi pada data pelatihan.

5. Evaluation dilakukan dengan menguji performa model pada data uji yang terpisah dari data latih. Evaluasi model didasarkan pada metrik akurasi, precision, recall, f1-score, dan confusion matrix. Analisis hasil evaluasi bertujuan untuk mengidentifikasi keunggulan dan kelemahan model dalam mengklasifikasikan setiap huruf hijaiyah.

$$Accuracy = (TP + TN) / (TP + FP + FN + TN)$$

$$Precision = TP / (TP + FP)$$

$$Recall = TP / (TP + FN)$$

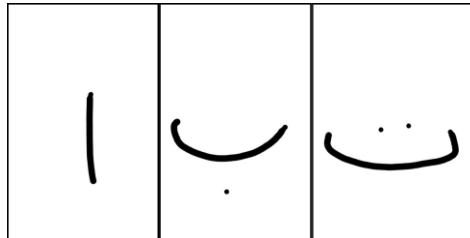
$$F1 - Score = 2(Precision Recall) / (Precision + Recall)$$

Keterangan: TP = True Positive, FP = False Positive, FN = False Negative

6. *Deployment*, adalah implementasi model pada dua platform aplikasi. Model hasil pelatihan diintegrasikan pada aplikasi web berbasis Flask sehingga dapat diakses melalui browser, serta pada aplikasi mobile yang dikembangkan menggunakan *React Native Expo*.

Pengguna, baik guru maupun siswa, dapat melakukan pengujian model dengan mengunggah gambar atau menggambar huruf hijaiyah secara langsung pada kedua aplikasi tersebut. Diharapkan, pendekatan multiplatform ini dapat memberikan manfaat lebih luas dan fleksibel dalam pembelajaran huruf hijaiyah di sekolah dasar.

HASIL DAN PEMBAHASAN



Gambar 3. Contoh Citra Huruf Hijaiyah dari Dataset Penelitian.

1. Dataset dan Pembagian Data

Penelitian ini menggunakan dataset citra huruf Hijaiyah yang telah melalui tahap pra-pemrosesan. Beberapa contoh citra dari dataset yang digunakan dapat dilihat pada **Gambar 3**, yang menunjukkan variasi tulisan tangan untuk huruf "Alif", "Ba", dan "ta". Total dataset yang digunakan terdiri dari 2100 citra, yang dibagi menjadi dua bagian utama untuk pelatihan dan pengujian model.

- Data Training: Sebanyak 1680 citra (80% dari total dataset) dimuat dari direktori dataset/train. Data ini digunakan secara eksklusif untuk melatih model Support Vector Machine (SVM).
- Data Testing: Sebanyak 420 citra (20% dari total dataset) dimuat dari direktori dataset/test_mixed. Data ini berfungsi sebagai data yang belum pernah dilihat sebelumnya (*unseen data*) untuk menguji performa model yang telah dilatih.

Setelah data dimuat, label kelas diubah menjadi format numerik melalui proses *label encoding* agar dapat diproses oleh algoritma.

2. Pencarian Hiperparameter (Hyperparameter Tuning)

Untuk mendapatkan performa model SVM yang optimal, dilakukan proses pencarian hiperparameter menggunakan metode GridSearchCV dengan validasi silang 5-fold (*5-fold cross-validation*). Proses ini secara sistematis menguji berbagai kombinasi parameter untuk menemukan konfigurasi terbaik. Parameter yang diuji mencakup C (parameter regularisasi), gamma (koefisien kernel), dan kernel (tipe kernel). Total terdapat 9 kandidat kombinasi parameter yang diuji, dan dengan 5-fold cross-validation, proses ini melibatkan total 45 kali proses *fitting* atau pelatihan model. Hasil dari proses tuning menunjukkan bahwa kombinasi parameter terbaik adalah sebagai berikut:

- Best Params: {'C': 10, 'gamma': 'scale', 'kernel': 'rbf'}
- Best CV Accuracy: 72.26%

Nilai akurasi validasi silang sebesar 72.26% ini merepresentasikan performa rata-rata model pada data training saat diuji dengan parameter terbaik.

3. Evaluasi Kinerja Model pada Data Uji

Setelah mendapatkan konfigurasi hiperparameter terbaik, model SVM dilatih kembali menggunakan seluruh data training (1680 citra) dan kemudian dievaluasi

performanya menggunakan 420 citra pada data testing.

Hasil evaluasi menunjukkan bahwa model mampu mencapai akurasi keseluruhan sebesar 83.57% pada data uji. Untuk mendapatkan wawasan yang lebih mendalam mengenai kinerja model pada setiap kelas huruf, dilakukan analisis menggunakan *Classification Report*

Tabel 1. Hasil Klasifikasi huruf hijaiyah

Huruf	Precision	Recall	F1-Score	Support
ain	0.80	0.86	0.83	14
alif	1.00	1.00	1.00	14
ba	0.70	1.00	0.82	14
dal	0.87	0.93	0.90	14
dhod	0.83	0.71	0.77	14
dzal	0.91	0.71	0.80	14
dzho	0.83	0.71	0.77	14
fa	0.62	0.71	0.67	14
ghoin	0.85	0.79	0.81	14
ha	1.00	1.00	1.00	14
ha'	0.77	0.71	0.74	14
hamzah	1.00	1.00	1.00	14
jim	1.00	1.00	1.00	14
kaf	1.00	1.00	1.00	14
kho	0.73	0.79	0.76	14
lam	1.00	1.00	1.00	14
lamalif	1.00	1.00	1.00	14
mim	1.00	1.00	1.00	14
nun	1.00	0.71	0.83	14
qof	0.50	0.50	0.50	14
ro	0.67	0.57	0.62	14
shod	0.80	0.86	0.83	14
sin	0.79	0.79	0.79	14
syin	0.80	0.86	0.83	14
ta	0.90	0.64	0.75	14
tho	0.75	0.86	0.80	14
tsa	0.75	0.86	0.80	14
wawu	0.80	0.86	0.83	14
ya	1.00	0.86	0.92	14
zain	0.65	0.79	0.71	14
accuracy			0.84	420
macro avg	0.84	0.84	0.84	420
weighted avg	0.84	0.84	0.84	420

Dari hasil klasifikasi tersebut, beberapa poin penting dapat dianalisis

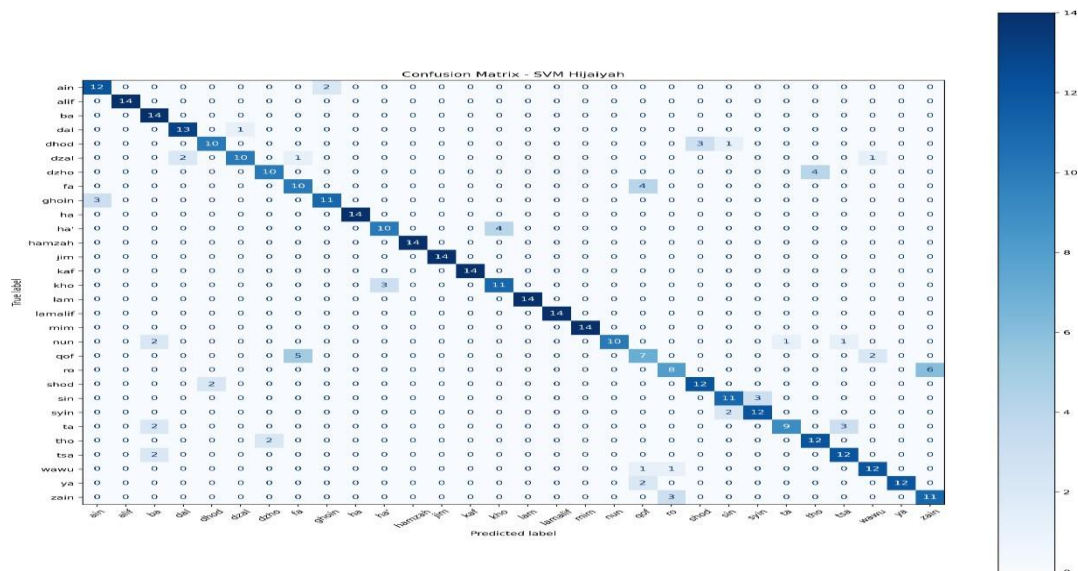
1. Performa Unggul: Model menunjukkan performa sempurna (precision, recall, dan F1- score bernilai 1.00) untuk beberapa huruf seperti 'alif', 'ha', 'hamzah', 'jim', 'kaf', 'lam', 'lamalif', dan 'mim'. Hal ini menandakan bahwa fitur-fitur dari huruf-huruf tersebut sangat distingtif dan mudah dikenali oleh model.
2. Performa Menengah: Sebagian besar huruf lainnya menunjukkan performa yang baik dengan F1-score di rentang 0.70 hingga 0.90.
3. Kelas yang lemah : Model mengalami kesulitan dalam membedakan beberapa huruf yang memiliki kemiripan visual. Huruf 'qof' menunjukkan F1-score terendah (0.50), diikuti oleh 'ro' (0.62) dan 'fa' (0.67). Rendahnya nilai ini kemungkinan disebabkan oleh kemiripan bentuk atau titik dengan huruf lain, yang menyebabkan model sering salah mengklasifikasikannya.

Secara keseluruhan, nilai rata-rata (macro dan weighted average) untuk precision, recall, dan F1-score mencapai 0.84, yang mengonfirmasi bahwa model memiliki kinerja yang solid dan seimbang di seluruh kelas, meskipun terdapat

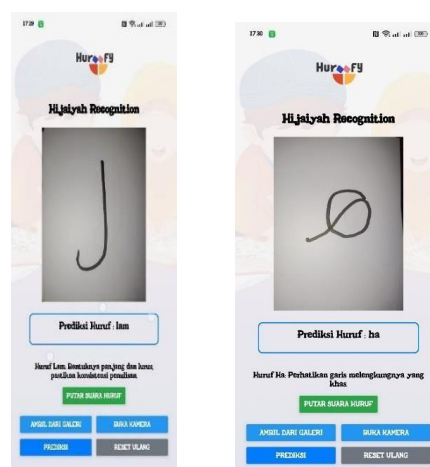
beberapa kelas yang sulit untuk dibedakan.

Analisis Confusion Matrix

Confusion matrix pada Gambar 3 menunjukkan hasil klasifikasi model SVM dengan kernel RBF ($C=10$, $\text{gamma}=\text{'scale'}$) terhadap data pengujian berupa 420 sampel tulisan tangan huruf hijaiyah anak usia 6–8 tahun. Matriks ini menampilkan jumlah prediksi yang benar dan salah untuk setiap kelas huruf hijaiyah. Nilai diagonal (warna biru tua) menunjukkan jumlah klasifikasi benar (True Positive), sedangkan sel lain di luar diagonal menunjukkan jumlah kesalahan klasifikasi atau salah prediksi (False Positive dan False Negative). Dari visualisasi tersebut terlihat bahwa beberapa huruf seperti "alif," "ha," "jim," "kaf," "lam," "lamalif," "mim," dan "hamzah" memiliki akurasi prediksi 100% karena bentuknya yang relatif konsisten dan mudah dibedakan. Sebaliknya, beberapa huruf seperti "qof," "ro," dan "fa" menunjukkan tingkat kesalahan yang relatif tinggi akibat bentuknya yang sering tertukar dengan huruf lain yang mirip. Secara keseluruhan, model SVM berhasil mencapai akurasi sebesar 83,57%, menunjukkan efektivitasnya dalam klasifikasi tulisan tangan huruf hijaiyah pada usia dini.



Gambar 4. Confusion Matrix



Gambar 5. Implementasi Antarmuka

Aplikasi Huroofy dirancang untuk memudahkan Anda mengenali tulisan tangan huruf hijaiyah secara intuitif. Cukup pilih gambar dari galeri dengan tombol AMBIL DARI GALERI atau ambil foto baru menggunakan BUKA KAMERA. Setelah gambar siap, tekan PREDIKSI untuk membiarkan sistem kecerdasan buatan menganalisis dan menampilkan hasilnya. Untuk melengkapi proses belajar, Anda bisa mendengarkan pelafalan yang benar dengan menekan PUTAR SUARA HURUF, dan jika ingin mencoba huruf baru, cukup tekan RESET ULANG untuk memulai kembali.

KESIMPULAN

Penelitian ini membuktikan bahwa algoritma *SVM* untuk klasifikasi tulisan tangan huruf hijaiyah pada anak usia 6-8 tahun. Model yang dibangun mencapai akurasi 83.57% dan dapat diintegrasikan pada aplikasi pembelajaran berbasis web/mobile. Pengembangan lebih lanjut dapat dilakukan dengan menambah variasi data dan menerapkan metode memperbanyak kreatif data agar memiliki akurasi model pengenalan huruf hijaiyah seperti *ensemble* atau algoritma *deep learning* lainnya untuk dipergunakan sebagai perbandingan model sehingga dapat meningkatkan akurasi yang signifikan untuk hasil lebih optimal.

DAFTAR REFERENSI

1. Arifin, B., A. Rosyid, dan R. Ramadhani. "Analisis Akurasi SVM pada Klasifikasi Tulisan Tangan Huruf Hijaiyah." *Jurnal Informatika dan Sains* 9, no. 2 (2023): 89–97.
2. Astuti, R. D., & Arifin, A. Z. "Komparasi Kinerja Algoritma Support Vector Machine dan Naive Bayes untuk Klasifikasi Tulisan Tangan Huruf Arab." *Indonesian Journal of Computing and Cybernetics Systems (IJCCS)*, 14, no. 3 (2020): 265-274.
3. Faturrahman, Irvan, dan Fitri Mintarsih Arini. "Pengenalan pola huruf hijaiyah khat kufi dengan metode deteksi tepi sobel berbasis jaringan syaraf tiruan backpropagation." *Jurnal Teknik Informatika* 11, no. 1 (2018): 37-46.
4. Hidayat, R., & Putra, A. B. W. "Pengembangan Aplikasi Media Pembelajaran Interaktif Pengenalan Huruf dan Angka untuk Anak Usia Dini Berbasis Android." *Jurnal Teknoinfo*, 16, no. 1 (2022): 123-129.
5. Marlina, S., & Suryadi, D. "Analisis Kesulitan Siswa Sekolah Dasar dalam Menulis Huruf Hijaiyah Bersambung." *Jurnal Basicedu*, 4, no. 4 (2020): 973-982.
6. Nasution, M. K., & Siregar, A. P. "Pengenalan Pola Huruf Hijaiyah Tunggal Menggunakan Ekstraksi Fitur Principal Component Analysis (PCA) dan Klasifikasi Support Vector Machine." *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6, no. 1 (2022): 359-370.
7. Nurdiansyah, D., A. Herlambang, dan P. I. Santosa. "Pengolahan Citra Digital untuk Pengenalan Karakter." *Jurnal Teknologi Informasi* 10, no. 2 (2021): 57–66.
8. Putra, R. P., dan N. Kusumawardani. "Ekstraksi Ciri HOG untuk Pengenalan Karakter Tulisan Tangan." *Jurnal Informatika: Jurnal Pengembangan IT* 7, no. 1 (2022): 23–30.
9. Rahman, A. U., & Abbas, N. "Handwritten Arabic Character Recognition using a Convolutional Neural Network." *International Journal of Advanced Computer*

- Science and Applications (IJACSA), 11, no. 12 (2020): 593-599.
10. Rosyid, A., dan B. Arifin. "Pengenal Karakter Huruf Hijaiyah Tulisan Tangan Menggunakan Support Vector Machine." *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)* 5, no. 3 (2021): 472–478.
 11. Sari, N. I., dan R. Ramadhani. "Penerapan HOG pada Pengenal Karakter Tulisan Tangan." *Jurnal Ilmiah Komputer dan Informatika* 8, no. 2 (2020): 123–128.
 12. Sihombing, R., dan R. Marpaung. "Analisis Evaluasi Model Machine Learning dengan Confusion Matrix." *Jurnal Teknologi Informasi* 16, no. 1 (2020): 12–18.
 13. Sulaiman, A., & Omar, K. "Comparative Study of Feature Extraction for Handwritten Arabic Character Recognition." *Journal of King Saud University - Computer and Information Sciences*, 33, no. 6 (2021): 694-702.
 14. Wahyuni, S. "Penerapan Media Visual dalam Pembelajaran Huruf Hijaiyah pada Anak Usia Dini." *Jurnal Pendidikan dan Konseling* 2, no. 2 (2020): 113–121.
 15. Yuliani, N., dan R. Surya. "Pembelajaran Interaktif untuk Meningkatkan Kemampuan Membaca Huruf Hijaiyah Anak Usia Dini." *Jurnal Pendidikan Anak Usia Dini* 3, no. 1 (2019): 10–17.