**Comparativa de enfoques y evolución del modelo**

**Objetivo y datos**

Mi objetivo fue construir una red neuronal capaz de clasificar **7 emociones** (angry, disgust, fear, happy, neutral, sad, surprise) a partir de imágenes **48×48** en escala de grises trabajando con diferentes conjuntos en tamaños de los splits de : **train ≈ 22.968**, **val ≈ 5.741** y **test ≈ 7.178** imágenes.

**Enfoques evaluados**

1. **MODELO 2 — CNN simple + augmentación**  
   Diseño secuencial básico (bloques Conv2D+MaxPool, Flatten y Dense con softmax). Apliqué augmentación moderada (rotación, traslación, flip) y *class weights* para desbalanceo.
2. **MODELO 6 — CNN compleja (LeakyReLU + BatchNorm)**  
   Aumenté capacidad y estabilidad: más capas convolucionales y densas, **BatchNormalization** tras conv, **LeakyReLU** para gradientes más estables, *early stopping* y *ReduceLROnPlateau*. Mantengo augmentación y pesos por clase.
3. **FINAL — CNN especializada + embeddings + router/ensamble**  
   Mantengo la CNN como **backbone** para extraer *embeddings* y añado dos capas de decisión:
   * Un **clasificador general** (softmax) sobre la salida de la CNN.
   * Un **clasificador lineal (LogisticRegression)** entrenado sobre *embeddings*; además, un **especializado** para el subgrupo *angry/fear/sad*.  
     Combino ambos con un **router por umbrales** y un pequeño **ensamble** (cuando el general está poco confiado, derivo al fino). En el general uso *label smoothing* para mejorar calibración.

**Evolución y verificación del enfoque final**

**A partir de la CNN entrenada (48×48 gris), evalué cuatro variantes en VAL y TEST:**

| **Variante (TEST)** | **Accuracy** | **Macro-F1** | **Comentario clave** |
| --- | --- | --- | --- |
| **Baseline — General SOLO** | **0,6311** | **0,5873** | **Softmax del backbone sin ayudas.** |
| **Especializado (TODAS LAS CLASES)** | **0,6434** | **0,6233** | **CNN como extractor + capa lineal; mejora notable en clases difíciles.** |
| **Router (General + Fino subset)** | **0,6311** | **0,5874** | **Cobertura baja (≈11,6% TEST); impacto global limitado.** |
| **Ensamble α=0,25** | **0,6454** | **0,6237** | **Mejor Macro-F1 y mejor accuracy global.** |
| **Ensamble α=0,50** | **0,6449** | **0,6193** | **Muy próximo al mejor; ligera caída en Macro-F1.** |
| **Ensamble α=0,75** | **0,6415** | **0,6055** | **Mezcla demasiado sesgada al general; pierde Macro-F1.** |

**El mejor compromiso en TEST lo obtuve con el Ensamble α=0,25 (Accuracy 0,6454, Macro-F1 0,6237). El Especializado puro también es fuerte (Macro-F1 0,6233), pero el pequeño ensamble aporta un plus consistente. El Router no mejora porque solo interviene en ≈9,9% (VAL) / 11,6% (TEST) de los casos.**

**Resumen en una frase: *Mi mejor configuración final es el Ensamble α=0,25, que supera al baseline y al router, y empata/prácticamente supera al especializado en Macro-F1 con la mejor accuracy global.***

**Evolucion de métricas por clase en los distintos modelos:**

* **fear: sube de 0,356 (Baseline) → 0,427 (Especializado) → 0,423 (Ensamble α=0,25).  
  *Es la clase más difícil; el especializado y el ensamble ayudan claramente.***
* **disgust: mejora grande de 0,482 (Baseline) → 0,617 (Especializado) → 0,611 (α=0,25).**
* **sad: de 0,490 → 0,523 (Especializado) → 0,526 (α=0,25).**
* **happy y surprise se mantienen altas (≈0,85 y ≈0,78 con α=0,25).**
* **neutral queda estable (≈0,60–0,61), con ligera mejoría en α=0,75 (0,614), a cambio de perder Macro-F1 global.**

**Conclusion: bloque difícil fear/sad/neutral es donde más gano con el enfoque especializado/ensamble, y disgust mejora de forma sustancial. Mantengo las clases “fáciles” (*happy/surprise*) sin sacrificar rendimiento.**

**Dónde se equivoca aún (a partir de matrices normalizadas)**

**Con Ensamble α=0,25 (TEST) se observa:**

* **fear → sad 20,4% y fear → angry 14,7%; fear → neutral 12,8%.**
* **sad → neutral 19,7% (y neutral → sad 15,3%).**
* **angry → neutral ~11–12% según variantes.**

**Esto confirma el solapamiento entre fear/sad/neutral/angry, ya visto en los cuadernos anteriores.**

**EVOLUCION DE MODELOS (MODELO 2 → MODELO 6 → FINAL)**

* **MODELO 2 (CNN simple): ~0,55 acc / 0,51 Macro-F1.**
* **MODELO 6 (CNN compleja, BN+LeakyReLU): ~0,61 / 0,58.**
* **FINAL (Ensamble α=0,25): 0,645 / 0,624.**

**Salto neto desde MODELO 6: +3,5 pts en accuracy y +4,4 pts en Macro-F1 aprox. La clave ha sido: usar la CNN como extractor de *embeddings*, sumar un clasificador lineal y combinar salidas (ensamble). El router confirma la idea de especialización, pero con poca cobertura su impacto final es menor.**

**Decisiones que justifican el resultado final**

* **Normalización interna (BN)** y **activaciones LeakyReLU/ReLU (según modelo)**: estabilidad y gradiente útil en capas profundas.
* **Augmentación y *class weights***: combaten desbalanceo y sobreajuste.
* **Label smoothing en el general → mejor calibración.**
* **LR sobre *embeddings* y submodelo fino para *angry/fear/sad* → más capacidad para fronteras difíciles.**
* **Ensamble ligero (α=0,25) → integra la confianza del general sin “tapar” al especializado; maximiza Macro-F1 y accuracy.**

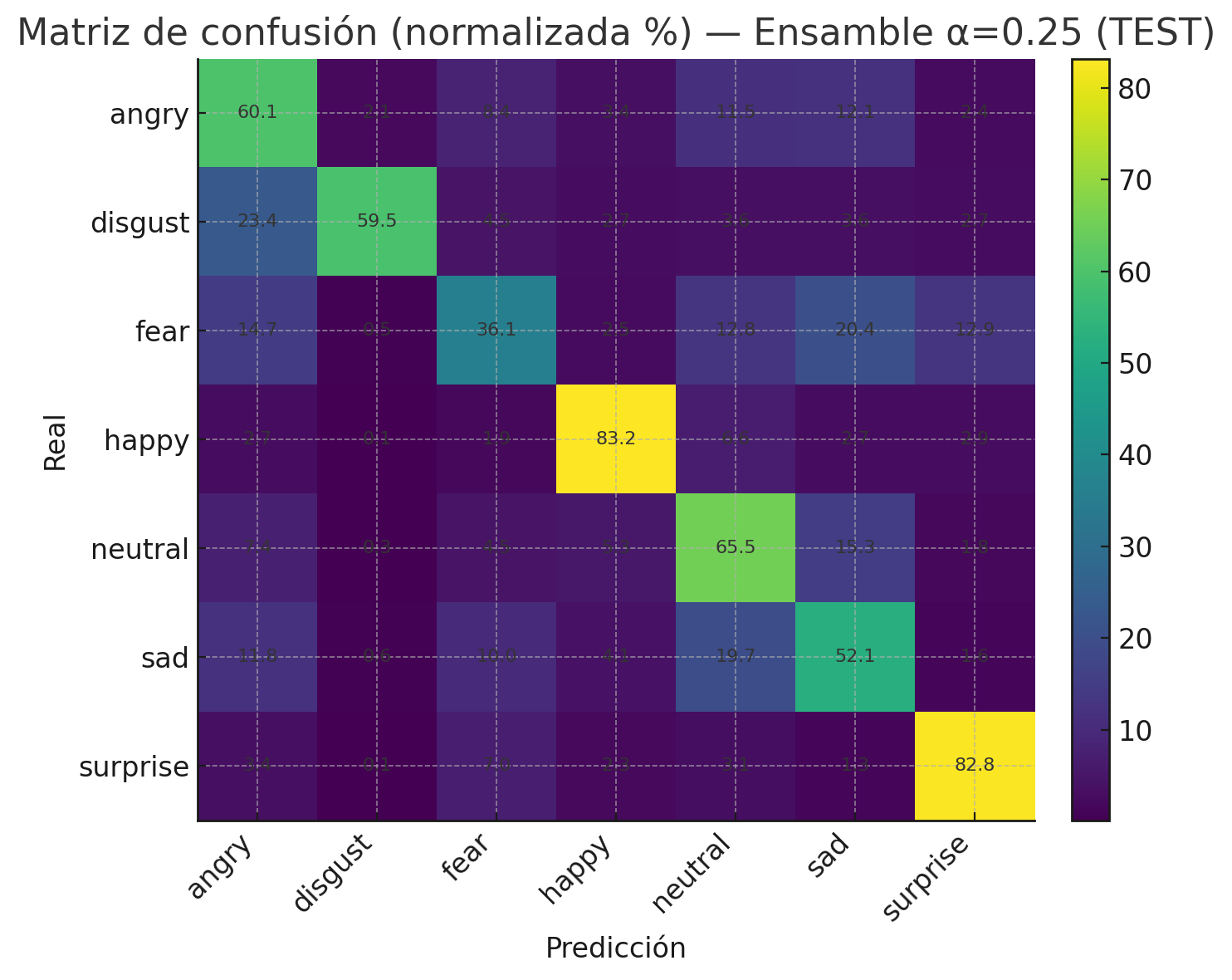
**6) Próximos pasos focalizados**

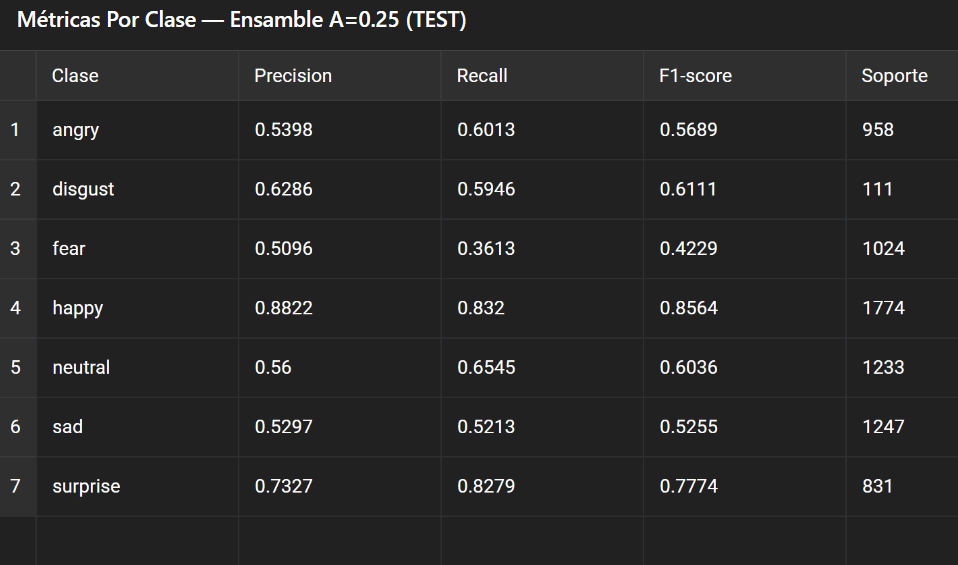
* **Optimizar α por Macro-F1 en VAL (búsqueda fina alrededor de 0,2–0,35).**
* **Aumentar señal en fear/sad: *oversampling* dirigido.**
* **Probar un modelo preentrenado como extractor (Mini-Xception/MobileNetV2 adaptado a 48×48) manteniendo la misma lógica de ensamble – no hay mucha seguridad porque William lo probo y no parecía bien**

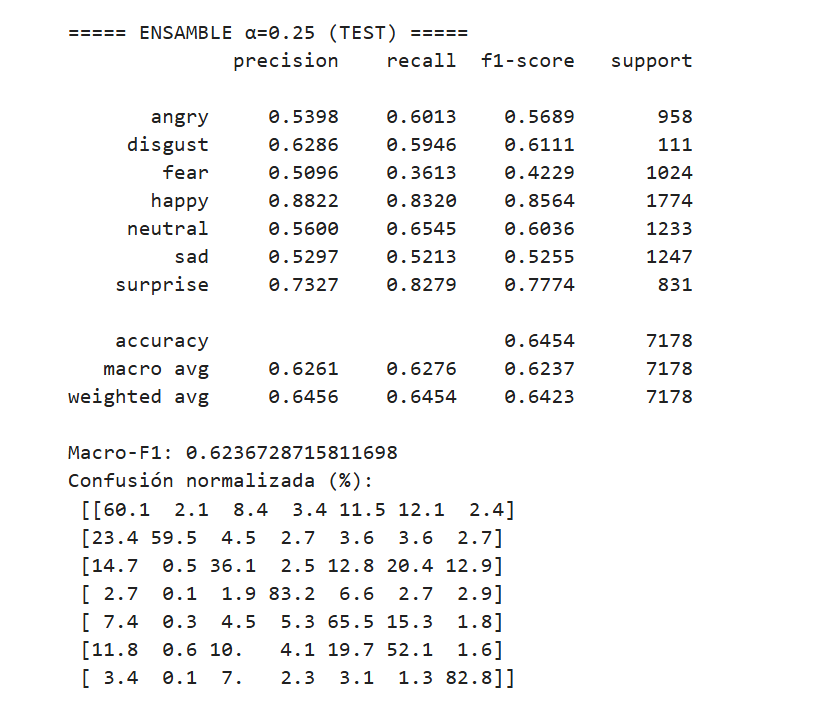
**CONCLUSION GENERAL**

**He validado que el mejor rendimiento global lo obtengo con un ensamble suave entre el clasificador general y el especializado (α=0,25), alcanzando 0,645 de accuracy y 0,624 de Macro-F1 en TEST. Este enfoque mejora de forma consistente a la línea base y al router, y mantiene el avance frente a mis modelos previos (MODELO 2 y MODELO 6). La ganancia se explica por:**

1. **extraer *embeddings* con la CNN:**
2. **añadir una capa lineal bien calibrada**
3. **combinar ambas salidas con un peso moderado. Los principales errores siguen concentrándose en *fear/sad/neutral* pero se confunden mucho menos, como se muestra en la matriz de confusión**

****

****

****